



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 171833

TO: Ginny Portner
Location: 3b02 / 3c18
Tuesday, November 29, 2005
Art Unit: 1645
Phone: 571-272-0862
Serial Number: 09 / 904994

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

STIC-Biotech/ChemLib

171833

From: Portner, Ginny
Sent: Wednesday, November 16, 2005 6:32 PM
To: STIC-Biotech/ChemLib
Subject: 09/904,994

please search SEQ ID NO 1, 2 and 3. thanks

Ginny Portner
Remsen Building
Art Unit 1645
Room E03, B02
(571) 272-0862

3C18

Scan into EDAN
RECEIVED
NOV 17 2005
STIC/CHEN. LIBRARY
(STIC)

Jan
22604

alrals



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor
22507, Remsen 1d86

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 – Circ. Desk



November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:29:19 ; Search time 14055 seconds
(without alignments)
11659.884 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgragattttccarctt.....aaaaagtagaacacagg 2883

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_st.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2880.6	99.9	2883	6	BD185302 Helicobac
2	2880.6	99.9	2883	6	AX356683 Sequence
3	2228	77.3	2452	6	BD185306 Helicobac
4	2228	77.3	2452	6	AX356695 Sequence
5	2158.6	74.9	2405	6	BD185303 Helicobac
6	2158.6	74.9	2405	6	AX356686 Sequence
7	2158.6	74.9	2407	6	BD185305 Helicobac
8	2158.6	74.9	2407	6	AX356692 Sequence
9	1948	67.6	2183	6	BD185304 Helicobac
10	1948	67.6	2183	6	AX356689 Sequence
11	1185.6	41.1	8406	1	AF330621 Helicobac
12	1138	39.5	2664	1	HECUREASE
13	1090.8	37.8	12037	1	AE001446
14	1081.6	37.5	10903	1	AE000529 Helicobac
15	1075.2	37.3	2735	6	AR054309 Sequence
16	1075.2	37.3	5100	1	HECUREABCD
17	1072.4	37.2	7088	1	AB032429 Helicobac
18	1062.6	36.9	2619	6	A41006 Sequence 1

19	1062.6	36.9	2619	6	A72509 Sequence 1
20	1062.6	36.9	2619	6	A73285 Sequence 1
21	1062.6	36.9	2619	6	AR061045 Sequence
22	1062.6	36.9	2619	6	AR161678 Sequence
23	1060.4	36.8	2430	1	AY227442 Helicobac
24	1058.6	36.7	4824	6	AR349305 Sequence
25	1058.6	36.7	4824	6	AR349306 Sequence
26	1057.8	36.7	2619	1	HPHUPABG
27	1056.8	36.7	2746	1	AF332656 Helicobac
28	1056.8	36.7	2746	6	A31515 H.pylori ur
29	985.2	34.2	6126	1	AF332656 Helicobac
30	985.2	34.2	30929	1	AE017145 Helicobac
31	894.6	31.0	1704	1	AY714224 Helicobac
32	885.8	30.7	1815	6	BD092530 Identific
33	884.8	30.7	1710	1	AY295085 Helicobac
34	881.8	30.6	1707	6	CQ903244 Sequence
35	881.8	30.6	1707	6	CQ903245 Sequence
36	880	30.5	1710	1	AF352376 Helicobac
37	878.4	30.5	1710	6	A07398 Helicobac
38	876	30.4	1680	1	AF479026 Helicobac
39	876	30.4	1680	1	BA000004_02
40	772.4	26.8	110000	1	BSUEROP_02
41	766	26.6	10118	1	BSUEROP_02
42	766	26.6	200690	1	BSUB0019
43	739.4	25.6	1665	1	AF508003 Helicobac
44	737.6	25.6	6537	1	BACUREA
45	733.8	25.5	4723	1	AB201709 Campyloba

ALIGNMENTS

RESULT 1	BD185302	BD185302	2883 bp	DNA	linear	PAT 17-JUN-2003
LOCUS	BD185302	Helicobacter felis vaccine.				
DEFINITION	BD185302	Helicobacter felis vaccine.				
ACCESSION	BD185302					
VERSION	BD185302.1	GI:31877502				
KEYWORDS	JP 2002355054-A/1.					
SOURCE	Helicobacter felis					
ORGANISM	Helicobacter felis					
REFERENCE	1 (bases 1 to 2883)					
AUTHORS	Kusters, J.G. and Cattoli, G.					
TITLE	Helicobacter felis vaccine					
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;					
COMMENT	AKZO NOBEL NV					
	OS Helicobacter felis					
	PN JP 2002355054-A/1					
	PD 10-DEC-2002					
	PF 16-JUL-2001 JP 2001214711					
	PR 17-JUL-2000 EP 00202565.8					
	PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI					
	PC					
	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC					
	A61K39/23,					
	PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,					
	PC C12N1/19,					
	PC					
	C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC					
	53,					
	PC					
	G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, PC					
	C12R1:01),					
	PC C12N15/00, C12N5/00, A61K37/02					
	CC Helicobacter felis vaccine					
	FH Key					
	FT CDS					
	FT CDS					
	Location/Qualifiers					
	(206) . . (886)					
	(897) . . (2603).					
	Location/Qualifiers					
	1..2883					
	/organism="Helicobacter felis"					

/mol_type="genomic DNA" /db_xref="taxon:214"									
ORIGIN									
Query Match 99.9%; Score 2880.6; DB 6; Length 2883; Best Local Similarity 100.0%; Pred. No. 0; Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	RGRAGATTTTCARCACTTCAAGCACATATTGATCCTGTGTTGTGGTGGTAAATTCR	60						
DB	1	RGRAGATTTTCARCACTTCAAGCACATATTGATCCTGTGTTGTGGTGGTAAATTCR	60						
QY	61	ACTTGTAAATCTATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA	120						
DB	61	ACTTGTAAATCTATTAATTAATTTTAAATAATTAATTAATTAATTAATTAATTA	120						
QY	121	TTACTATATTAATAAGTTAATAAAGTAACGAATTAGGACTAATAATCCATGCCCT	180						
DB	121	TTACTATATTAATAAAGTTAATAAAGTAACGAATTAGGACTAATAATCCATGCCCT	180						
QY	181	TAAAAATTTAACAAAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
DB	181	TAAAAATTTAACAAAGAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAAGTTCTT	240						
QY	241	GTATATTAATGCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA	300						
DB	241	GTATATTAATGCGGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA	300						
QY	301	ACCGNAGCATTTGCTTACATTAAGTCCCATATTAATGGACGAGCGCGGTGGAAAAAA	360						
DB	301	ACCGNAGCATTTGCTTACATTAAGTCCCATATTAATGGACGAGCGCGGTGGAAAAAA	360						
QY	361	AACCGTTGCCAGCTTATGGAAGAGTGCACTTTTGAAGAAAAGATGAAGTAAATGCC	420						
DB	361	AACCGTTGCCAGCTTATGGAAGAGTGCACTTTTGAAGAAAAGATGAAGTAAATGCC	420						
QY	421	CGGGTGGGTAAATATGGTTCCCGATCTAGGTGTAGAAGCCACCTTTCCTGATGGTAGAA	480						
DB	421	CGGGTGGGTAAATATGGTTCCCGATCTAGGTGTAGAAGCCACCTTTCCTGATGGTAGAA	480						
QY	481	ACTTGTAACTGTGAATTTGGCCCATCGAAACGAGATGAGCACTTCAAAGCGGGCGAAGTGAA	540						
DB	481	ACTTGTAACTGTGAATTTGGCCCATCGAAACGAGATGAGCACTTCAAAGCGGGCGAAGTGAA	540						
QY	541	ATTTGGTTCGATAAAGACATCAGAGCTCAATGCAGGCAAGAGTAACCCAACTTGAGGT	600						
DB	541	ATTTGGTTCGATAAAGACATCAGAGCTCAATGCAGGCAAGAGTAACCCAACTTGAGGT	600						
QY	601	TACTAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAA	660						
DB	601	TACTAATGAAGGGCCCTAAATCTTGCATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAA	660						
QY	661	CAAGGCACCTAAATTCGATCGTGAAGAAAGCCTATGGCAAAACGCTAGATATTCCTCTGG	720						
DB	661	CAAGGCACCTAAATTCGATCGTGAAGAAAGCCTATGGCAAAACGCTAGATATTCCTCTGG	720						
QY	721	CACACGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG	780						
DB	721	CACACGCTACGATTTGGGGCAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGG	780						
QY	781	CAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATTAACATCGCGGATGAACGCCATAA	840						
DB	781	CAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATTAACATCGCGGATGAACGCCATAA	840						
QY	841	ACATAAGCGCTTGACAAGCGCAAAATCTCAACGATTTATCAAGTAAGGAGACTCCCATGA	900						
DB	841	ACATAAGCGCTTGACAAGCGCAAAATCTCAACGATTTATCAAGTAAGGAGACTCCCATGA	900						
QY	901	AAATGAAAAAACAAGATATGTAAATACCTACCGAACCAACCAAGGGCGATAAAGTGGCT	960						
DB	901	AAATGAAAAAACAAGATATGTAAATACCTACCGAACCAACCAAGGGCGATAAAGTGGCT	960						
QY	961	TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACCACTATGGCGAAGAAC	1020						

DB	961	TAGGAGATACCGATCTTTGGGCAGAGTAGAACAATGACTATACCACTATGGCGAAGAAC	1020						
QY	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGTATGGGTGAGAGCAATAGCCCTGATG	1080						
DB	1021	TTAAATTTGGCGGGTAAAACTATCCGTGAGGTATGGGTGAGAGCAATAGCCCTGATG	1080						
QY	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGATGATTATCGATACACCGGATTTACA	1140						
DB	1081	AAAAACCCCTAGATTTAGTTCATCACTAACCGATGATTATCGATACACCGGATTTACA	1140						
QY	1141	AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGCAAGAAACAGG	1200						
DB	1141	AAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGCATTTGGCAAGCAAGAAACAGG	1200						
QY	1201	ACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGGGCACAGAACACTAGCAG	1260						
DB	1201	ACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGGGCACAGAACACTAGCAG	1260						
QY	1261	GGGAAGGTATGATTTATTACCGCTGGGGGAATCGATTCAACACCCCACTTCCTTTCTCCAC	1320						
DB	1261	GGGAAGGTATGATTTATTACCGCTGGGGGAATCGATTCAACACCCCACTTCCTTTCTCCAC	1320						
QY	1321	AACAATTTCCCTACCGCTTAGCCCAATGGCGTTACAAACCATGTTGGAGGGCGCACAGGTC	1380						
DB	1321	AACAATTTCCCTACCGCTTAGCCCAATGGCGTTACAAACCATGTTGGAGGGCGCACAGGTC	1380						
QY	1381	CTGTAGATGACCAATGCGACTACTATCACTCCGGGCAATGGACCTTGACCCGATGT	1440						
DB	1381	CTGTAGATGACCAATGCGACTACTATCACTCCGGGCAATGGACCTTGACCCGATGT	1440						
QY	1441	TGCGCGCAGCAGAGAGTATTTCTATGATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
DB	1441	TGCGCGCAGCAGAGAGTATTTCTATGATGTGGGCTTTTGGGCAAAAGCAATAGCTCTA	1500						
QY	1501	GCAAAAACAACCTTGTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAG	1560						
DB	1501	GCAAAAACAACCTTGTAGAACAGTAGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAG	1560						
QY	1561	ACTGGGGCAACAACCAAGTGCATCGATCACTCTTGGAGCGTGCGACATGAATACGATG	1620						
DB	1561	ACTGGGGCAACAACCAAGTGCATCGATCACTCTTGGAGCGTGCGACATGAATACGATG	1620						
QY	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGGTTATGATGACACCCCTAA	1680						
DB	1621	TGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCAAGGTTATGATGACACCCCTAA	1680						
QY	1681	ATGCAATGAACGGCGGCCATCCATGCCCTACCATTTGAGGGAGCGGGTGGAGGACACT	1740						
DB	1681	ATGCAATGAACGGCGGCCATCCATGCCCTACCATTTGAGGGAGCGGGTGGAGGACACT	1740						
QY	1741	CACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
DB	1741	CACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCA	1800						
QY	1801	CTATTTCCCTATACCATTAATACGGTTGCAGAACCTTAGACATGCTCATGACATGCCACC	1860						
DB	1801	CTATTTCCCTATACCATTAATACGGTTGCAGAACCTTAGACATGCTCATGACATGCCACC	1860						
QY	1861	ACCTTAGACAAAACGATCCGCGAGGATTTACAATTTTCTCAAAAGCGGTATCCCGCCGGCT	1920						
DB	1861	ACCTTAGACAAAACGATCCGCGAGGATTTACAATTTTCTCAAAAGCGGTATCCCGCCGGCT	1920						
QY	1921	CTATCGCGGTGAAGATGTGCTCCATGATATGGGTGTGATCCGATGACAAAGCTCGGAT	1980						
DB	1921	CTATCGCGGTGAAGATGTGCTCCATGATATGGGTGTGATCCGATGACAAAGCTCGGAT	1980						
QY	1981	CGCAAGCAATGGGGCTGCAGCGAAGTGTCTCTCGAACTTTGGCAGACTGCGGATAAGA	2040						
DB	1981	CGCAAGCAATGGGGCTGCAGCGAAGTGTCTCTCGAACTTTGGCAGACTGCGGATAAGA	2040						
QY	2041	ATAAAAAAGAAATTTGGTAAGCTTCTGGAAGATGGCAAGATAACGATTAATTTCCGATTA	2100						
DB	2041	ATAAAAAAGAAATTTGGTAAGCTTCTGGAAGATGGCAAGATAACGATTAATTTCCGATTA	2100						

QY 481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
DB 481 ACTTGTAACTGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAA 540
QY 541 ATTGGTTCCGATAAAGACATCGAGCTCAATGAGGCAAGAGTAACCGAACTTGAGGT 600
DB 541 ATTGGTTCCGATAAAGACATCGAGCTCAATGAGGCAAGAGTAACCGAACTTGAGGT 600
QY 601 TACTAATGAAGGCGCTAAATTCCTTGTGATGTGGGTAGCCATTTTCCACTTCTTTTGAAGCTAA 660
DB 601 TACTAATGAAGGCGCTAAATTCCTTGTGATGTGGGTAGCCATTTTCCACTTCTTTGAAGCTAA 660
QY 661 CAAGGCACTAAATTCGATCGTGAAGGCAAGGCTATGAGCAAGGCTTAGATATTCCTCTCTGG 720
DB 661 CAAGGCACTAAATTCGATCGTGAAGGCAAGGCTATGAGCAAGGCTTAGATATTCCTCTCTGG 720
QY 721 CAACAGCTACGATTTGGGCGAGGCAAAACCCGCAAGGTGCAAGTTCCTCTCTGGTGG 780
DB 721 CAACAGCTACGATTTGGGCGAGGCAAAACCCGCAAGGTGCAAGTTCCTCTCTGGTGG 780
QY 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACGCCATAA 840
DB 781 CAGTAAAAAGTGAATTTGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACGCCATAA 840
QY 841 ACATAAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTTCAAGTAAGGAGACTCCCATGA 900
DB 841 ACATAAAGCGCTTGAACAAGCGGAAATCTCAAGGATTTTCAAGTAAGGAGACTCCCATGA 900
QY 901 AAATGAAAAAACAAGTAATGTAAATACCTACGGACCCACCAAGGCGGATAAAGTGGCT 960
DB 901 AAATGAAAAAACAAGTAATGTAAATACCTACGGACCCACCAAGGCGGATAAAGTGGCT 960
QY 961 TAGGAGATACCGATCTTTGGGCAAGAGTAGAAATGACTATACCATCTATGGCGAAGAAC 1020
DB 961 TAGGAGATACCGATCTTTGGGCAAGAGTAGAAATGACTATACCATCTATGGCGAAGAAC 1020
QY 1021 TTAATTTTGGCGGGTAAATATCATCGTGAAGGTATGGGTACAGAAATAGCCCTGATG 1080
DB 1021 TTAATTTTGGCGGGTAAATATCATCGTGAAGGTATGGGTACAGAAATAGCCCTGATG 1080
QY 1081 AAAACACCTAGATTTAGTATCATCACTAAACCGGATGATTTAGTATGATGATGATTTTACA 1140
DB 1081 AAAACACCTAGATTTAGTATCATCACTAAACCGGATGATTTAGTATGATGATGATTTTACA 1140
QY 1141 AAGCCGACATTTGGGATTTAAACCGGCAAAATCCATGGCATTTGGCATTTGGCATTTGGCAT 1200
DB 1141 AAGCCGACATTTGGGATTTAAACCGGCAAAATCCATGGCATTTGGCATTTGGCATTTGGCAT 1200
QY 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGTTGGGCACAGAGCACTAGCAG 1260
DB 1201 ACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGTTGGGCACAGAGCACTAGCAG 1260
QY 1261 GGAAGGTATGATTTACCGCTTGGGGGAATCGAATTCACACCCACTTCTTTTCTCCAC 1320
DB 1261 GGAAGGTATGATTTACCGCTTGGGGGAATCGAATTCACACCCACTTCTTTTCTCCAC 1320
QY 1321 AACAAATTCCTACCGCTTAGCAATGCGGTTCACACCAATGTTTGGAGGCGGCACAGGTC 1380
DB 1321 AACAAATTCCTACCGCTTAGCAATGCGGTTCACACCAATGTTTGGAGGCGGCACAGGTC 1380
QY 1381 CTGTAGATGGCAGGAATGCGACTACTATCACTCCGGGCAAAATGGAATTTGACCGGATGT 1440
DB 1381 CTGTAGATGGCAGGAATGCGACTACTATCACTCCGGGCAAAATGGAATTTGACCGGATGT 1440
QY 1441 TGGCGCAGCAGAAAGATATTTATGAAATGTTGGGCTTTTGGGCAAGGCAATAGCTCTTA 1500
DB 1441 TGGCGCAGCAGAAAGATATTTATGAAATGTTGGGCTTTTGGGCAAGGCAATAGCTCTTA 1500
QY 1501 GCNAAAAAAACAATTTAGAACAAAGTAGAAGCGGCGGATTTGGTTTAAATTTGATGAAG 1560
DB 1501 GCNAAAAAAACAATTTAGAACAAAGTAGAAGCGGCGGATTTGGTTTAAATTTGATGAAG 1560
QY 1561 ACTGGGGCACAAACACCAAGTGGCATCGATCACTGCTTGGCGGTGGCAGATGAATACGATG 1620

DB 1561 ACTGGGGCACAAACACCAAGTGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATG 1620
QY 1621 TGCAAGTTTCTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAA 1680
DB 1621 TGCAAGTTTCTATCCACACCGATACAGTCAATGAGGCGAGGTTATGTAGATGACACCCCTAA 1680
QY 1681 ATGCAATGAACGGGCGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 1740
DB 1681 ATGCAATGAACGGGCGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 1740
QY 1741 CACTGATGTTATCACCATGCGAGGCGAGCTCAATATTTCTACCTCTCTCCACACCCCA 1800
DB 1741 CACTGATGTTATCACCATGCGAGGCGAGCTCAATATTTCTACCTCTCTCCACACCCCA 1800
QY 1801 CTATTCCTTATACCATTAATACGGTTGAGAACTTATAGATGCTCTATGATGCTATGATGCTATGAT 1860
DB 1801 CTATTCCTTATACCATTAATACGGTTGAGAACTTATAGATGCTCTATGATGCTATGATGCTATGAT 1860
QY 1861 ACCTAGACAAACGATCCGAGGATTTACAAATTTTCTCAAAGCGTATCCGCCCGGCT 1920
DB 1861 ACCTAGACAAACGATCCGAGGATTTACAAATTTTCTCAAAGCGTATCCGCCCGGCT 1920
QY 1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGATCGCGATGACAAAGCTCGGATTT 1980
DB 1921 CTATCGCGCTGAAGATGTGCTCATGATATGGGTGATCGCGATGACAAAGCTCGGATTT 1980
QY 1981 CGCAAGCAATGGGCGGTGCGAGGCGAAGTATTCCTCGAACTTTGGCAGACTCGGATGA 2040
DB 1981 CGCAAGCAATGGGCGGTGCGAGGCGAAGTATTCCTCGAACTTTGGCAGACTCGGATGA 2040
QY 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAGATAAAGATAAATTTCCGCAATTA 2100
DB 2041 ATAAAAAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAGATAAAGATAAATTTCCGCAATTA 2100
QY 2101 AGCGTACATCTCCAAATACATATCAACCCCGCTTTGAACCCAGCGGTGAGCGAGTATA 2160
DB 2101 AGCGTACATCTCCAAATACATATCAACCCCGCTTTGAACCCAGCGGTGAGCGAGTATA 2160
QY 2161 TCGGCTCTGTGGAGAGGCGAAGATCGCGACTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2220
DB 2161 TCGGCTCTGTGGAGAGGCGAAGATCGCGACTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2220
QY 2221 GCGTAAAAACCAAAATCGTATCAAGCGGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2280
DB 2221 GCGTAAAAACCAAAATCGTATCAAGCGGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2280
QY 2281 CTAAACCGCTCTGTGCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCATCAG 2340
DB 2281 CTAAACCGCTCTGTGCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCATCAG 2340
QY 2341 GCAAGGCGAAATTTGACACCGAGTCACTTTGTTTCCAAAGTCCCTATGAAAATGGCG 2400
DB 2341 GCAAGGCGAAATTTGACACCGAGTCACTTTGTTTCCAAAGTCCCTATGAAAATGGCG 2400
QY 2401 TGAAGAAAAAGCTGCGGCTTAGAGCGCAAGTTCCTACCGGTCAAAACTGCGGTAAACATCA 2460
DB 2401 TGAAGAAAAAGCTGCGGCTTAGAGCGCAAGTTCCTACCGGTCAAAACTGCGGTAAACATCA 2460
QY 2461 CCAAGAAAGACTTCAAGTTTCAACCGCAAAACCGGCAAAATCAACCGTTCGATCCGAAAACCT 2520
DB 2461 CCAAGAAAGACTTCAAGTTTCAACCGCAAAACCGGCAAAATCAACCGTTCGATCCGAAAACCT 2520
QY 2521 TCAGAGTCTTTGTAGATGGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
DB 2521 TCAGAGTCTTTGTAGATGGGCAAACTCTGCACTCTAAACCCACCTCGCAAGTGCCTCTAG 2580
QY 2581 CCAGCGCTACATTTCTTCTAGGCAAACTGCCCCCTTTGGGGCAGGTTATTTTAGGAA 2640
DB 2581 CCAGCGCTACATTTCTTCTAGGCAAACTGCCCCCTTTGGGGCAGGTTATTTTAGGAA 2640
QY 2641 TCTTCATCAACACGACCTGCAATCGGTCTTTCGCTGTGCGATCGGTCTTTCGCTTTTAAACCAAC 2700

Db 2641 TCTTCATCAACGCACTCGTCAATCGGTCTTCGCGTGTGCGATCGTGTGCGCTTTAAACAAC 2700
Qy 2701 TTTTCATCTTTAAAGCAATCGCCATTTTTTAATTAATTAATTTAAATCTTTATTAATTAATATAT 2760
Db 2701 TTTTCATCTTTAAAGCAATCGCCATTTTTTAATTAATTAATTTAAATCTTTATTAATTAATATAT 2760
Qy 2761 TATGCCCTCATTTTTTAAAGGAGAATATGCGTAGGTCTTTGGTATTTGCTATGCTGCGGG 2820
Db 2761 TATGCCCTCATTTTTTAAAGGAGAATATGCGTAGGTCTTTGGTATTTGCTATGCTGCGGG 2820
Qy 2821 TTGCTTTGGTCTGGCGCAAGAGGTATTGAAACCCATCGCTCAAAAAGGTAGAGGCCAC 2880
Db 2821 TTGCTTTGGTCTGGCGCAAGAGGTATTGAAACCCATCGCTCAAAAAGGTAGAGGCCAC 2880
Qy 2881 AGG 2883
Db 2881 AGG 2883

RESULT 3
BD185306
LOCUS BD185306 2452 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185306
VERSION BD185306.1 GI:31877506
KEYWORDS JP 2002355054-A/5.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2452)
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 5 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/5
PD 10-DEC-2002
PP 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
53,
PC G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC
C12R1/01),
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (48)..(728)
FT CDS (739)..(2445).

FEATURES
source 1..2452
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
Qy 159 AGGACTAATATCCATTCGCTTTAAATTTAAACAAGGAGTAAATAGGTGAATCTCAC 218
Db 1 AGGACTAATATCCATTCGCTTTAAATTTAAACAAGGAGTAAATAGGTGAATCTCAC 60
Qy 219 CCAAGAGCAGAAAAGTCTTGTATATTATGCGGCGAAGTGGCTAGAAAGCGCAAG 278

Db 61 CCAAAGAGCAAGAAAAGTTCTTTATATTATTCGGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGTGCCCATATATATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGTGCCCATATATATGG 180
Qy 339 ACGAAGCGCGCTGGGAAAAAAACCGTTGCCAGCTTATGGAAGAGTGCATGCATTTT 398
Db 181 ACGAAGCGCGCTGGGAAAAAAACCGTTGCCAGCTTATGGAAGAGTGTATGSCATTTT 240
Qy 399 TGAAGAAAGATGAAGTAAATGCGCGGGTGGGTAAATATGTTTCCCGATCTAGTGTAGAAG 458
Db 241 TGAAGAAAGATGAAGTAAATGCGCGGGTGGGTAAATATGTTTCCCGATCTAGTGTAGAAG 300
Qy 459 CCACCTTTCTCTGATGGTACGAAACTTTGTAATCTGTAACTTGGCCCATCGAACCCAGATGAGC 518
Db 301 CCACCTTTCCCGATGGCACCAACTCTGTAATCTGTAATGGCCCATCGAACCTGATGAAC 360
Qy 519 ACTTCAAGCGCGGCGAAGTGAAATTTGGTTCGATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTTAAGCGGGTGAAGTGAATTTGGCTGTGATAAAGACATTGAACTCAACGCGAGTA 420
Qy 579 AAGAAGTAAACCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 638
Db 421 AGGAAGTTTACCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 480
Qy 639 ATTTCCACTCTTTGAAGCTAACAGGCACATAAATTCGATCGTGAAGAAAGCTTATGGCA 698
Db 481 ATTTCCACTCTTTGAAGCCAAAGACATTTGAATTCGATCGGAAAGAAAGCTTATGGCA 540
Qy 699 AACGCTAGATATTTCCCTCTGCGCAACACGCTACGCAATGGGGCAGGACAAACCCGCAAG 758
Db 541 AACGCTAGATATTTCCCTCTGCGCAACACATACGCAATGGGGCAGGACAAACCCGTAAG 600
Qy 759 TGCAGTTGATTCCTTTGGTGGCGAGTAAAAAGTGAATGGGCAATGAACGGGCTTGTGAATA 818
Db 601 TGCAGTTAAATCCCTCTTGGCGGTAGTAAAAAGTGAATGGGCAATGAACGGGCTTGTGAATA 660
Qy 819 ACATCGCGGATGAAGCCGCAATGAAGCGCTTGACAAGCGCAATCTACGGATTTA 878
Db 661 ATATTTGGCGGACGACGCGCCATAACCAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
Qy 879 TCAAGTAGGAGAGCTCCCATGAAAATGAAAAAACAAGAAATATGTAATACTACCGACCC 938
Db 721 TCAAGTAGGAGAGCTCCCATGAAAATGAAAACAAGAGATATGTAATACTACCGACCC 780
Qy 939 ACCAAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACATGAC 998
Db 781 ACCACAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCGAGAGTAGAACATGAC 840
Qy 999 TATACCACTATGGGGAAGAACTTAAATTTGGCGGGGTAAACATCTCCGTGAGGATATG 1058
Db 841 TATACCACTATGGGGAAGAACTTAAATTTGGCGGGGTAAACATCTCCGTGAGGATATG 900
Qy 1059 GGTCAAGCAATAGCCCTGATGAAAACACCCCTAGATTTAGTTCATCACTAACCCGATGAT 1118
Db 901 GGTCAAGCAATAGCCCTGATGAAAACACCCCTAGATTTAGTTCATCACTAACCCGATGAT 960
Qy 1119 ATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAGCGCAAAATTCATGGC 1178
Db 961 ATTGACTACACCGGGATTTTACAAAGCCGACATTTGGCAATTAAGATGGCAAAATTCATGGC 1020
Qy 1179 ATTGGCAAGCGAAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1238
Db 1021 ATTGGCAAGCGAAGAAACAGGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGGT 1080
Qy 1239 GTGGGCACAGAGACATAGCAGGGGAGGTATGATTTATCCGCTGGGGGATCCGATTCA 1298
Db 1081 GTGGGCACAGAGACATAGCAGGGGAGGTATGATTTATCCGCTGGGGGATCCGATTCA 1140
Qy 1299 CACACCCACTTCTCTTCTCCCAACAATTCCTACCGCTCTAGGCAATGGCGCTTACCAAC 1358

```
Db 1141 CACACCCACTTCTCTCCACAAAGAAATCCCTACCGCTCTAGCCAAATGGCGTTACAA 1200
QY 1359 ATGTTTGGAGCGGCGACAGGTCTGTAGATGGCAGAAATGCGACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGCAGACGCCCGGTAGATGGCAGAAATGCGACTACTATCACTCCGGGC 1260
QY 1419 AAATGGAATCTGCACCGCATGTTGCGCGAGCAGAAAGAGTATTTCTATGAATGTGGCTTT 1478
Db 1261 AAATGGAATCTGCACCGCATGTTGCGCGAGCAGAAAGAGTATTTCTATGAATGTGGCTTT 1320
QY 1479 TTGGGCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGAACAAAGTAGAGCGGCGCG 1538
Db 1321 TTGGGCAAGGCAATAGCTCTAGTAAAAAACAACTTGTAGAACAAAGTAGAGCGGCGCG 1380
QY 1539 ATTGGTTTTAAATTCGATGAAGACTGGGCGCAACACCAAGTCGATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTTAAATTCGATGAAGACTGGGCGCAACAACTCCAAGTCGATCGATCACTGCTTG 1440
QY 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTAGCAGATGAATACGATGTGCAAGTTTGTATACACACCGATACAGTCAATGAGGCA 1500
QY 1659 GGTATGTAGATGACACCCCTAAATGCAATGAAGCGGCGGCCATCCATGCCTACCAATT 1718
Db 1501 GGTATGTAGATGACACCCCTAAATGCAATGAAGCGGCGGCCATCCATGCCTACCAATT 1560
QY 1719 GAGGAGCGGCTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCAATATT 1778
Db 1561 GAGGAGCGGCTGGAGGACACTCACCTGATGTTATCACCATGCGAGCGAGCTCAATATT 1620
QY 1779 CTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAACGTTTGCAGAACCTTA 1838
Db 1621 CTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAACGTTTGCAGAACCTTA 1680
QY 1839 GACATGCTCATGACATGCCACCACTAGACAAACGATCGCGAGGATTTACATTTTCT 1898
Db 1681 GACATGCTTATGACCTGCCACCACTAGATAAACGATCGCGAGGATTTCCAAATTTCT 1740
QY 1899 CAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATGGGTGTG 1958
Db 1741 CAAAGCGGTATCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATGATATCGGTGTG 1800
QY 1959 ATCGGATGACAGCTCGAATTCGCAAGCAATGGGCGGTGCAGCGCAAGTATTCTCGA 2018
Db 1801 ATCGGATGACAGTTCCGATTCGCAAGCAATGGGCGGTGCAGCGCAAGTATTCTCTAGA 1860
QY 2019 ACTTGCAGATCGCGATGAAGATAAAGAAATTTGGTAAGCTTCTCTGAAGATGGCAAA 2078
Db 1861 ACTTGCAGAACTGCAGACAGAAATAAAGAAATTTGGTAAGCTTCTCTGAAGATGGTGA 1920
QY 2079 GATAAGATAATTTCCGCATTAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTCG 2138
Db 1921 GATAATGACAACTTCCGCATCAACAGCTATATCTCCAAATACACTATTAATCCCGCTTTCG 1980
QY 2139 ACCACGGGCTGAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAGATCGCGCACTTGGTG 2198
Db 1981 ACCATGCGCTGAGCGAGTATATCGGCTCTGTGGAAAGAGGCGCAAGATCGCGCACTTGGTG 2040
QY 2199 GTGTGAATCTCGCTCTTTTGGCGTAAACCCAAATCTGTGATCAAAAGCGGTATGGTG 2258
Db 2041 GTGTGAATCTCGCTCTTTTGGCGTAAACCCAAATCTGTGATCAAAAGCGGTATGGTG 2100
QY 2259 GTCTTCTCTGAATGGCGATTTCTAAGCGTCTGTGCCACTCTCCCAACCGGTTATTAC 2318
Db 2101 GTGTCTCTGAATGGCGATTTCTAATGCGTCTGTGCCACTCTCTCAGCGGTTATTAC 2160
QY 2319 CGCGAAATGTTTGGGCATACCGCAAGGCGAAATTTGACACCAAGCATCACTTTTGTTC 2378
Db 2161 CGCGAAATGTTTGGGCATACCGCAAGGCGAAATTTGACACCAAGCATCACTTTTGTTC 2220
QY 2379 AAAGTCGCTATGAAATGCGGTGAAGAAAGCTGGCTTAGAGCGCCAAAGTTCTACCG 2438
Db 2221 AAAGTCGCTATGAAATGCGGTGAAGAAAGAACTAGGTTTAGAGCGCAAGGTGCTCCCC 2280
```

```
QY 2439 GTCAAAAACCTGCCGTAAACATCAACAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAA 2498
Db 2281 GTGAAAAACCTGCCGTAAACATCAACAAGAAAGACTTCAAGTTCAACGACAAACCGCAAAA 2340
QY 2499 ATCACCGTCGATCGAAACCTTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCACCGTCGATCGAAACCTTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCCACCTCGCAAGTGCCTCTAGCCACGCTTACACTTTCTTCTAGGCACAAT 2610
Db 2401 CCCACCTCGCAAGTGCCTCTAGCCACGCTTACACTTTCTTCTAGGCAATAAT 2452
```

RESULT 4

```
AX356695
LOCUS AX356695
DEFINITION Sequence 13 from Patent EP1176192.
ACCESSION AX356695
VERSION AX356695.1 GI:18674032
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
```

REFERENCE

```
1. Kusters, J.G. and Cattoli, G.
AUTHORS Helicobacter felis vaccine
TITLE Patent: EP 1176192-A 13 30-JAN-2002;
JOURNAL Akzo Nobel N.V. (NL)
```

FEATURES

```
source
1..2452
Location/Qualifiers
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
48..728
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23263.1"
/db_xref="GI:18674033"
/transl_table=11
/translation="MKLTPKEQKFLYYAGVARKKAEGLKLNQNPALAIYISAHIM
DEARRKKTVAELMECEMFLKKDEVNPGVMVPLDGEATPDGKTLKLVNWNPIEP
DEHFKAEGVFQCDKDIELNAGKEVTELVNPGKSLHVGSHFPAFKALKFDE
KATGKRLDIPSGNTRLRIGAGTRKVLQILPGSKKVIQMGNLVNNIADERHKHKALD
AKSHGFLK"
```

CDS

```
739..2445
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/db_xref="GI:18674034"
/transl_table=11
/translation="MRMKQEVNNTYPTGDKVRLGDTDLWAEVEHDYTYTGBELKF
GAKTIREGMGQNSPDENTLDLVTNAMIIDYGIYKADIGIKNGKIHGIGKNGKD
MQQGVPHMVVGVGTALAGEGMIITAGGIDSHFLSPQQPFTALANGVTMTFGGTT
GPVDGNTATTITGKWNLRMLRAAEVSMNVGFLGKNSSSKKQLVQVQVGAIGFK
LHEDMCTPSAIDHCLSVADBYDVQVCIHTDTVNEAGYVDVDTLNMNGRAIHAVHLE
AGGHSPPDVIIMAGEVNLPSSTPTTPIYINTVAEHLDMLTCHLDRKIRIDELQFS
QSRIRPGSIAEDVLHDIGVIAMTSSDSQAMGRAGEVIRPTWQTADKNKKGFKLQPD
GANDNFRIKRYISKYITNPALTVYSGVEGKIADLVNWNPAFPGVKPKTVIK
GMVVFSEMGDSNASVTPQPVYVREMFHGHGKAFDTSITFVSKVAYENGVKELGL
ERKVLPVKNCRNITKKDFKNDKTAKITVDPKTFEVFDGKLTCKTSKPTSEVPLAQRYT
FF"
```

CDS

```
739..2445
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/db_xref="GI:18674034"
/transl_table=11
/translation="MRMKQEVNNTYPTGDKVRLGDTDLWAEVEHDYTYTGBELKF
GAKTIREGMGQNSPDENTLDLVTNAMIIDYGIYKADIGIKNGKIHGIGKNGKD
MQQGVPHMVVGVGTALAGEGMIITAGGIDSHFLSPQQPFTALANGVTMTFGGTT
GPVDGNTATTITGKWNLRMLRAAEVSMNVGFLGKNSSSKKQLVQVQVGAIGFK
LHEDMCTPSAIDHCLSVADBYDVQVCIHTDTVNEAGYVDVDTLNMNGRAIHAVHLE
AGGHSPPDVIIMAGEVNLPSSTPTTPIYINTVAEHLDMLTCHLDRKIRIDELQFS
QSRIRPGSIAEDVLHDIGVIAMTSSDSQAMGRAGEVIRPTWQTADKNKKGFKLQPD
GANDNFRIKRYISKYITNPALTVYSGVEGKIADLVNWNPAFPGVKPKTVIK
GMVVFSEMGDSNASVTPQPVYVREMFHGHGKAFDTSITFVSKVAYENGVKELGL
ERKVLPVKNCRNITKKDFKNDKTAKITVDPKTFEVFDGKLTCKTSKPTSEVPLAQRYT
FF"
```

ORIGIN

```
Query Match 77.3%; Score 2228; DB 6; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 159 AGGACTATAATCCCATTCCTTTAAATTTAAACAAGAGTAAATAGGTGAAACTCACAC 218
Db 1 AGGACTATAATCCCATTCCTTTAAATTTAAACAAGAGTAAATAGGTGAAACTCACAC 60
```


Qy 219 CCAAAGAGCGAAGAAAGTTCTTGTATATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 278
Db 61 CCAAAGAGCGAAGAAAGTTCTTGTATATATATATCGGGCGAAGTGGCTAGAAAGCGCAAG 120
Qy 279 CAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTCCCATATTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCTTACATTAGTCCCATATTATGG 180
Qy 339 ACGAAGCGCGCTGGGAAAAAACCCTTGCAGCTTATGGAAGAGTGCATGCACTTTT 398
Db 181 ACGAAGCGCGCTGGGAAAAAACCCTTGCAGCTTATGGAAGAGTGTATGCACTTTT 240
Qy 399 TGA AAAAGATGAAGTAAATGCCCCGGGTGGGTAAATATGGTCCCCGATCTAGGTGTAGAG 458
Db 241 TGA AAAAGACGAGGTGATGCCCGGGTGGGGAATATGGTCCCTGATTTGGGCGTGGAG 300
Qy 459 CCACCTTCTCTGATGGTACGAACCTTGTAACTGTGAATTTGGCCCATCGAACCATGAGC 518
Db 301 CCACCTTCCCCTGATGGCACCAACTCGTAACTGTGAATTTGGCCCATCGAACCTGATGAAC 360
Qy 519 ACTTCAAAGCGGGCGAAGTGAATTTGCTTGCATATAAGACATCGAGCTCAATGCAGCA 578
Db 361 ACTTAAAGCGGGTGAAGTGAATTTGCTGTGATAAGACATTTGAACCTCAACGAGSTA 420
Qy 579 AAGAAGTAAACCGAATTTGAGGTTACTAATGAAGGGCCCTAAATCCTTGCACTGGGTAGCC 638
Db 421 AGGAAGTTACCGAATTTGAGGTTACTAATGAAGGGCCCTAAATCCTTGCACTGGGTAGCC 480
Qy 639 ATTTCCACTTTTGAAGCTTAAAGGCACTTAAATTCGATCGTGTAAAGACCTTATGGCA 698
Db 481 ATTTCCACTTTTGAAGCAACAAAGCAATTTGAATTTGATCGTGGGCAAAAGCCCTATGGCA 540
Qy 699 AAGCCTAGATATTCCTCTGGCAACACGCTACGATTTGGGCGAGGCAACCCGCAAG 758
Db 541 AAGCCTAGATATTCCTCTGGCAACACACATGACATTTGGGCGAGGCAAAAGCCCTATGGCA 600
Qy 759 TGCAGTTGATTCCTCTTGGTGGAGTAAAAAGTGAATTTGGCATGAAACGGGCTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAACGGGCTTGTGAATA 660
Qy 819 ACATCGCGGATGAACGCCATTAACATAAAGCGCTTGAACGGCGAAATCTCACGGATTTA 878
Db 661 ATATTGGCGAGCAAGCCATAACACAAAGCGCTAGACAAAGCAAAATCTCACGGATTTA 720
Qy 879 TCAAGTAAGAGACTCCCATGAAAAATGAAAAACAAAGATATGTAATATCTACGAGACC 938
Db 721 TCAAGTAAGAGACTCCCATGAAAAATGAAAAACAAAGATATGTAATATCTACGAGACC 780
Qy 939 ACCAAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 998
Db 781 ACCAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGAC 840
Qy 999 TATACCACTTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATG 1058
Db 841 TATACCACTTATGGCGAAGAACTCAATTCGGTGCAGGTAAAACTATCCGTGAGGGTATG 900
Qy 1059 GTTCAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCACTAACCGCATGATT 1118
Db 901 GGTTCAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCACTAACCGCATGATT 960
Qy 1119 ATCGACTACACGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAAATCCATGGC 1178
Db 961 ATTGACTACACGGGATTTACAAAGCCGACATTTGGCATTAAAAATGGCAAAATCCATGGC 1020
Qy 1179 ATTGGCAAGGCGAGAAACAAAGGACATGCAAGATGGCGTGAAGCCCTCATATGGTCTGGGT 1238
Db 1021 ATTGGCAAGGCGAGAAACAAAGGACATGCAAGATGGCGTGAAGCCCTCATATGGTCTGGGT 1080
Qy 1239 GTGGGCACAGAACCTAGCAGGGGAAGGTATGATTATACCGCTGGGGGAATCGATTCA 1298
Db 1081 GTGGGCACAGAACCTAGCAGGGGAAGGTATGATTATACCGCTGGGGGAATCGATTCA 1140
Qy 1299 CACACCCACTTCTTCTCCACAACAAATTCCTCCCTACCGCTCTAGCCCAATGGCGTTACAAAC 1358

Db 1141 CACACCCACTTCTCTCTCCACAACAAATTCCTTACCGCTCTAGCCCAATGGCGTTTACAACA 1200
Qy 1359 ATGTTTGGAGGGCGGCACAGGTCTCTGTAGATGCGACGAATGCGCACTACTATCACTCCGGGC 1418
Db 1201 ATGTTTGGCGGTGGACAGGCCCTGTAGATGCGACGAATGCGCACTACTATCACTCCGGGC 1260
Qy 1419 AAATGGAACTTGCACCGCATGTTGGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1478
Db 1261 AAATGGAACTTGCACCGCATGTTGGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTT 1320
Qy 1479 TTGGCAAAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGG 1538
Db 1321 TTGGCAAAAGGCAATAGCTCTAGTAAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGG 1380
Qy 1539 ATTGGTTTAAATTCGATGAAGCTTGGGGCAACAACAAGTGCATCGATCACTGCTTG 1598
Db 1381 ATTGGTTTAAATTCGATGAAGCTTGGGGCAACAATTCGAAGTGCATCGATCACTGCTTG 1440
Qy 1599 AGCGTGGCAGATGAATACGATGTGCAAGTTTCTATCCACACGATACAGTCAATGAGGCA 1658
Db 1441 AGCGTGGCAGATGAATACGATGTGCAAGTTTCTATCCACACGATACAGTCAATGAGGCA 1500
Qy 1659 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCATGCTTACCACTT 1718
Db 1501 GGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCATGCTTACCACTT 1560
Qy 1719 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGCGAGGCGAGTCAATATT 1778
Db 1561 GAGGAGCGGTGGAGGACACTCACTGATGTTATCACCATGCGAGGCGAGTCAATATT 1620
Qy 1779 CTACCTCTCCACACCCCTACTATCCCTATACCATTAATACCGTTCGAGAACACTTTA 1838
Db 1621 CTACCTCTCCACACCCCTACTATCCCTATACCATTAATACCGTTCGAGAACACTTTA 1680
Qy 1839 GACATGCTCATGACATGCGCACCCCTAGACAAACGATCCGCGAGGATTTACAAATTTCT 1898
Db 1681 GACATGCTCATGACATGCGCACCCCTAGATTAACGATCCGCGAGGATCTCCAATTTCT 1740
Qy 1899 CAAAGCGGTATCCGCGCGCTCTATCGGGCTGAAGATGTGCTCATGATATGGGTG 1958
Db 1741 CAAAGCGGTATCCGCGCGCTCTATCGCGCTGAAGATGTGCTCATGATATCGGTG 1800
Qy 1959 ATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGAGGCGAAGTTCCTCGA 2018
Db 1801 ATCGGATGACAAAGTTCGCAAGCAATGGGCGGTGAGGCGAAGTTCCTCGA 1860
Qy 2019 ACTTGGCAGACTCGGATTAAGAAATAAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAA 2078
Db 1861 ACTTGGCAAACTGACAGACAGATTAAGAAATTTGGTAAAGCTTCTGAAAGATGCAAA 1920
Qy 2079 GATAACGATAATTTCCGATTAAGCGCTACATCTCAAAATACATATCAACCCCGCTTTG 2138
Db 1921 GATAATGACAACTTCGATCAAAAGCTATATCTCCAAATACACCAATTAATCCCGCTTTG 1980
Qy 2139 ACCACGCGGTGAGGAGTATATCGGCTCTGTGGAAGGGCAAGATCCCGACTTTGGT 2198
Db 1981 ACCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAGGGCAAGATCCCGACTTTGGT 2040
Qy 2199 GTGTGGAATCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAAGCGGTATGGTG 2258
Db 2041 GTGTGGAATCTGCTTTTGGCGTAAAAACCAAAATCGTGATCAAGCGGTATGGTG 2100
Qy 2259 GTCTTCTGAAATGGGCGATTCTAACCGCTCTGTGCCACTCCCCAACCGGTTTATTAC 2318
Db 2101 GTGTTCTGAAATGGGCGATTCTAAATGGCTGTGTGCCACTCTCAGCGGTTTATTAC 2160
Qy 2319 CGCAAAATGTTTGGGCACTACGCGAAGCGGAAATTTGACACGAGCATCACTTTGTTTC 2378
Db 2161 CGCAAAATGTTTGGGCACTACGCGAAGCGGAAATTTGACACGAGCATCACTTTGTTTC 2220
Qy 2379 AAAGTCCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAGCGCCCAAGTTCTACCG 2438

Db 2221 AAAGTCGCTATGAAAATGGTGTGAAGAAAACCTAGGTTTAGAGCGCAAGGTGCTCCCC 2280
QY 2439 GTCAAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2498
Db 2281 GTGAAAACCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAACGGCAAAA 2340
QY 2499 ATCAGCGTCGATCCGAAAACCTTCGAGGTCCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2558
Db 2341 ATCAGCGTCGATCCGAAAACCTTCGAGGTCCTTTGTAGATGGCAAACTCTGCACCTCTAAA 2400
QY 2559 CCCACTGCAAGTCGCTCTAGCCAGCGCTACACTTCTCTAGGCACAAT 2610
Db 2401 CCCACTGCAAGTCGCTCTAGCCAGCGCTACACTTCTCTAGGCACAAT 2452

RESULT 5
LOCUS BD185303 2405 bp DNA linear PAT 17-JUN-2003
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185303
VERSION BD185303.1 GI:31877503
KEYWORDS JP 2002355054-A/2.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2405)
AUTHORS Kusters, J.G. and Cattoli, G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: JP 2002355054-A 2 10-DEC-2002;
AKZO NOBEL NV
COMMENT OS Helicobacter felis
PN JP 2002355054-A/2
PD 10-DEC-2002
PF 16-JUN-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI
PC C12N15/09, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC
A61K39/23
PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,
PC C12N1/19,
PC C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC
53,
PC G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1.01), (C12Q1/68, PC
C12R1.01),
PC C12N15/00, C12N5/00, A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (1)..(681)
FT CDS (692)..(2398).
FEATURES
source Location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"

ORIGIN
Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 206 GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTTATATATGCGGCGAAGTGGCT 265
Db 1 GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTTATATATGCGGCGAAGTGGCT 60
QY 266 AGAAGGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325
Db 61 AGAAGGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGAAGCCATTGCTTACATTAGT 120
QY 326 GCCCATATTATGACGAGCGCGCTGGGAAAAAACCCTGTGCCACGCTTATGGAAGAG 385

Db 121 GCCCATATTATGACGAGCGCGCTGGGAAAAAACCCTGTGTAACCTTATGGAAGAA 180
QY 386 TGCATGCACATTTTGAATAAAGATGAAGTAATGCCCGGGTGGGTAAATATGGTTCCTCCGAT 445
Db 181 TGTATGCACTTTTGAATAAAGATGAGTGTATGCCCGGTGGGGAATATGGTCCCTGAT 240
QY 446 CTAGGTGTAGAAGCCACCTTTCCCTGTAGTGTGTAAGAACTTGTAACTGTGAAATTTGGCCCATC 505
Db 241 TTGGGCGTAGAAGCCACTTTCCCGATGGGCAACCACTCGTAAACCGTGAATTTGGGCCCAT 300
QY 506 GAACCCAGATGAGCACTTCAAGCGGGGGAAGTGAATTTGGTTGCGATTAAGACATCGAG 565
Db 301 GAACCTGATGAACACACTTTAAAGCGCGGTGAAGTGAATTTGGCTGTGTATTAAGACATTGAG 360
QY 566 CTCAATGCAAGCAAGAAAGTAAACCGAACTTCAGGTGTACTAATGAAGGCGCTAAATCCCTTG 625
Db 361 CTCAACGCGGTGAAGAGTTACGAGCTTGAAGTTACCAAGAAAGGACCTTAATTCCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACAGGCACTAAAAATTCGATCGTGA 685
Db 421 CATGTGGGTAGCCATTTCCACTTCTTTGAACCAACCAAGGCAATTGAAATTCGATCGGGA 480
QY 686 AAAGCTATGCAAAACGCTTAGATATTCCCTCTGGCAACACGCTACGATTTGGGCGAGA 745
Db 481 AAAGCTATGCAAAACGCTTAGATATTCCCTCTGGCAACACGCTACGATTTGGGCGAGA 540
QY 746 CAAACCCGCAAGTGCAAGTTGATTTCTTCTGGTGCAGTAAAGAAAGTGTATTCGATGAAC 805
Db 541 CAAACCCGTAAGTGCAAGTTAATCCCTCTTGGCGGTAGTAAAGAAAGTGTATTCGATGAAC 600
QY 806 GGGCTTGTGAATAACATCGCGGATGAACGCACTAAACATAAAGCGCTTGACAAGGCGAAA 865
Db 601 GGGCTTGTGAATAATTTGCGACGAAGCCCTAAACACAAAGCACTAGACAAGGCAAAA 660
QY 866 TCTCAGCGATTTATCAAGTAAGAGAGCTCCCATGAAATGAAAAAACAAGAAATATGTAAA 925
Db 661 TCTCAGCGATTTATCAAGTAAGAGAGCTCCCATGAAATGAAAAAACAAGAGTATGTAAA 720
QY 926 TACCTACGAGCCCAACCAAGGCGATTAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAGA 985
Db 721 CACCTACGAGCCCAACCAAGGCGATTAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAGA 780
QY 986 AGTAGAACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
Db 781 AGTAGAACATGACTATACCACTATGGCGAAGAGCTCAATTTGGCGGGGTAAAACTAT 840
QY 1046 CCGTAGGCGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCATCAC 1105
Db 841 CCGTAGGCGGTATGGGTGAGAGCAATAGCCCAAGATGAAAAACACCTTAGATTTAGTCATCAC 900
QY 1106 TAAGCGGATGATTTAGTACACTACACCGGATTTCAAGCGGACATTTGGATTTAAAAACGG 1165
Db 901 CAACCGGATGATTTAGTACACTACACCGGATTTTATAAGCGGACATTTGGTATTTAAAAATGG 960
QY 1166 CAAATCCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGTAAAGCCTCA 1225
Db 961 CAAATCCATGGCATTTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGTAAAGCCTCA 1020
QY 1226 TATGCTGCTGGGTGAGGACAGAGCACTAGCAGGGAAGGTATGATTTATACCGCTCG 1285
Db 1021 TATGCTGCTGGGTGAGGACAGAGCACTAGCAGGGAAGGTATGATTTATACCGCTCG 1080
QY 1286 GGGAAATCGATTCACACCCACTTCTCTTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1345
Db 1081 GGGGATCGATTCGACACCCACTTCTCTCTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1140
QY 1346 TGGGTTACACCATGTTTGGAGGCGGACAGGTCCTGTAGATGACCAAGTGCAGAACTGCAGTAC 1405
Db 1141 TGGGTTTACACCATGTTTGGAGTGGGACAGGTCCTGGTAGATGGCAGCAATTCGGACAC 1200
QY 1406 TATCACTCCGCGGCAAAATGGAACCTTGCACCGCATTTTGGCGGACAGAGAGTATTCTAT 1465


```
Db 1201 CATCACTCGGGGAAATGGAATTGTCACCGCATGTTGCGCGCAGCTGGAAGAGTATTCTAT 1260
Qy 1466 GAATGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTGTAGAAACAAGT 1525
Db 1261 GAATGGGCTTTTGGGCAAGGCAATAGCTCTAGTAAAAAACAACCTGTAGAACAGT 1320
Qy 1526 AGAAGCGGCGGATTTGGTTTAAATTTGANTGAAGACTGGGGCACAACACCAAGTGGCAT 1585
Db 1321 AGAAGCGGCGGATTTGGTTTAAATTTGANTGAAGACTGGGGCACAACACCAAGTGGCAT 1380
Qy 1586 CGATCACTGCTTGAGCGTGGCAGATCAATACGATGTGCAAGTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGAGCGTGGCAGATCAATACGATGTGCAAGTTGTATCCACACCGATAC 1440
Qy 1646 AGTCAATGAGCGGATTTAGTAGATCACACCTTAAATGCAATGAACGGGCGCGCATCCA 1705
Db 1441 GGTCAATGAGCGGATTTAGTAGATCACACCTTAAATGCAATGAACGGGCGCGCATCCA 1500
Qy 1706 TGCCTACCACTTGAAGGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1765
Db 1501 TGCCTACCACTTGAAGGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGG 1560
Qy 1766 CGAGTCAATATTTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGTCAATATTTACCTCTCTCCACACCCCACTATTCCCTATACCAATTAATACGGT 1620
Qy 1826 TGCAGAACCTTAGACATGCTATGATGCGACACCACTAGACAAACGATCCGCGAGGA 1885
Db 1621 TGCAGAACCTTAGACATGCTATGATGCGACACCACTAGACAAACGATCCGCGAGGA 1680
Qy 1886 TTTTACAATTTTCTCAAGCGGTATCGCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTCCCAAGCGGTATCGCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA 1740
Qy 1946 TGATATGGGTGTGATCGGATGAACAGCTCGGATTCGCAAGCAATGGGCGTCAAGCGGA 2005
Db 1741 TGATATGGGTGTGATCGGATGAACAGCTCGGATTCGCAAGCAATGGGCGTCAAGCGGA 1800
Qy 2006 AGTGATTCCTCGAATTTGGCAGACTCGGATGAAGATGAAGATTAAGATTTGGTAACTTCC 2065
Db 1801 AGTGATTCCTCGAATTTGGCAGACTCGGATGAAGATGAAGATTAAGATTTGGTAACTTCC 1860
Qy 2066 TGAAGATGCAAGATGAATTTTCGCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1861 TGAAGATGTCAGATGAATGAACATTCGCGATCAACCGCTATATCTCCAAATACACTAT 1920
Qy 2126 CAACCCCGCTTTGACCCACCGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGT 2185
Db 1921 TAAATCCCGCTTTGACCCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGT 1980
Qy 2186 CGCGACCTTGGGTGGAAATCTCGCTTTTGGCGGTAAACCCCAAAATCGTGATCAA 2245
Db 1981 CGCGACCTTGGGTGGAAATCTCGCTTTTGGGTGTAACCAACCCCAAAATCGTGATCAA 2040
Qy 2246 AGGCGGTATGTTGCTCTCTGAATGGCGATTTCAACGCTCTGTGCGCACTCCCCA 2305
Db 2041 AGGCGGTATGTTGCTCTCTGAATGGCGATTTCAACGCTCTGTGCGCACTCCCCA 2100
Qy 2306 ACCGGTTTATACCGGAAATGTTTGGGCATCACGGCAAGCGGCAAAATTTTGACACCGCAT 2365
Db 2101 GCGGTTTATACCGGAAATGTTTGGGCATCACGGCAAGCGGCAAAATTTTGACACCGCAT 2160
Qy 2366 CACTTTGTTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAGCTGGCTTAGAGCG 2425
Db 2161 CACTTTGTTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAAATCGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAAGTGTACCGGTGAAAAATCGCGGTAAACATCACTAAGAAAGACTTCAAAATTCACAA 2280
Qy 2486 CAAAACGGCAAAAATCACCGTGCATCCGAAAACCTTCGAGGTCCTTGTAGATGGCAACT 2545
Db 2281 CAAAGCGGCGCATATCACTGTGATCTTAAACCTTCGAGGTCCTTGTAGATGGCAACT 2340
```

```
Qy 2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTCTTCTAGGC 2605
Db 2341 CTGCACCTCTAAACCCCGCTCTGAAGTGCCTCTAGCCCAACGCTACACTTCTTCTAGGC 2400
Qy 2506 ACAAT 2610
Db 2401 ACAAT 2405
RESULT 6
LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 4 from Patent EP1176192.
ACCESSION AX356686
VERSION AX356686.1 GI:18674023
KEYWORDS Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
AUTHORS Kusters,J.G. and Cattoli,G.
TITLE Helicobacter felis vaccine
JOURNAL Patent: Ep 1176192-A 4 30-JAN-2002;
Akzo Nobel N.V. (NL)
FEATURES
source Location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="unassigned DNA"
/db_xref="taxon:214"
1..681
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23257.1"
/db_xref="GI:18674024"
/translation="MKLTPKOEKFLLYAGEVARKKAEGLKNQPEALAYISAHIM
DEARRGKTVAELEMECHFLKDEVMGVGNMVPDLGVEATFPDGTKLVTNWPPEP
DEHFAGEVFGCDKDIENAGKEVTELEVTNEGPKSLHVGSHFFHFFETNKALKKDRE
KAYGRUDIPSGNTRLRIGAGQTRKVLPIPLGGSKVKVIGMNLVNNIADIERHKKHALDK
AKSHGFYK"
692..2398
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23258.1"
/db_xref="GI:18674025"
/translation="MKMKQEVNVTYGTGDKVRLGDTDLMAEVEHDYTTYGEELKF
GAKTIREGMGQNSPDENTLDLVTNAMIIDYTIYKADIGIKNGKIHGIGKAGNKD
MQGVSHMVVVGCTEALAGEMLITAGGIDSHFLPSQOFPPTALANGVTVMFGGCT
GPVDGNATITPGKWNLMRLAAEYSMMVFLGKGNSSSKQLVQVEAGALGFK
LHEDWGTSPSIDCHLSVADEYDVQVCITHTDVNEAGYVDDTLNMGRAITHAYHIEG
AGGSHPDVITMAGELNLPSSSTPTPTINTVAEHLDMMLTCHLLDKR IREDLQFS
QSRIRPGSIAAEDVLHDIGVITAMTSSDSQAMGRAGEVTPRTWOTADKNKKEPKLPEP
GADNDFRIKRYISKYTNINPALTHGVSEVIGSVEGKIADLVVNPAPFVGKPKIVIK
GMVDFSEMGDSNASVTTPQVYTYREMFHGHGKAKPDTISITFVSKVAEVENGVEKLGL
ERKVLVKNCRNITTKDKFKNNKTAHTITVDPKTFEVDGKLCITSKSPASEVPLAQRYT
FF"
```

ORIGIN

```
Query Match 74.9%; Score 2158.6; DB 6; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
Qy 206 GTGAACTCACACCAAGAGCAAGAAAGTCTTGTATATATATGCGGCGCAAGTGGCT 265
Db 1 GTGAACTCACACCAAGAGCAAGAAAGTCTTGTATATATGCGGCGCAAGTGGCT 60
Qy 266 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAGGCATTGCTTTACATTAGT 325
Db 61 AGAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCCGAGGCATTGCTTACATTAGT 120
```

QY	326	GCCCATATTATGACGAAAGCGCGCGTGGAAAAAACCCTGTCACGCTTATGGAAGAG	Db	1201	CATCACTCCGGGCAAAATGGAACTTGCCCGCATGTTGGCGCAGCTGAAGAGTATTCTAT	1260
Db	121	GGCCCATATTATGACGAGCGCGCGTGGCAAAAAACCCTTGCTGAACTTATGGAAGAA	QY	1466	GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAAACAACCTTGTAGAACAAAGT	1525
QY	386	TGCATGCACTTTTGAANAAGTGNAGTAATGCCCCGGGGTGGTAAATATGTTCCCGAT	Db	1261	GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCAGTAAAAAAACAACCTGTAGAACAAAGT	1320
Db	181	TGTATGCACTTTTGAANAAGATGAGGTGATGCCCGGTGTGGGAATATGTTCCCTGAT	QY	1526	AGAAGCGGGCGCGATTGGTTTTAAATTTGCAAGACTTGGGGGCAACAACAAGTGCAGT	1585
QY	446	CTAGGTGTAGAAAGCCACCTTCTCTGATGGTACGAAACTTGTAACTGTGAATTTGGCCCATC	Db	1321	AGAAGCGGGCGCGATTGGCTTTAAATTTGCATGAAGACTGGGGGCAACAACAAGTGCAGT	1380
Db	241	TTGGGCGGTAGAGCCACTTTCCTCCGATGGCACCAAACTCGTAACCGTGAATTTGGCCCAT	QY	1586	CGATCACTGCTTGAGCGGTGGCAGATGAATAGATGTGCAAGTTTGTATTCACACCGCATAC	1645
QY	506	GAAACAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG	Db	1381	CGATCACTGCTTGAGCGGTAGCAGATGAATAGATGTGCAAGTTTGTATTCACACCGCATAC	1440
Db	301	GAAACCTGATGAACACTTTAAAGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTGAG	QY	1646	AGTCAATGAGCGCAGTTATCTAGATGACACCCATAATGCAATGAACGGCGCGCCATCCA	1705
QY	566	CTCAATGCAAGGCAAGAAGTAACCGAACTTGTAGGTTACTTAATGAAGGGCCTAATCCTTG	Db	1441	GGTCAATGAGCGCAGTTATCTAGATGACACCCATAATGCGATGAACGGCGCGCCATCCA	1500
Db	361	CTCAACGCGGGTAAGGAAGTTACCGAGCTTGAAGTTTACCACGAAGGACCTAAATCCTTG	QY	1706	TGCTTACCACTTGAGGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGG	1765
QY	626	CATGTGGGTAGCCATTTCACATTCTTTGAAAGCTTAAAGGCATTAATAATTCGATCGTGA	Db	1501	TGCTTACCACTTGAGGGAGCGGGCGAGGACACTCACCTGATGTTATCACCATGGCAGG	1560
Db	421	CATGTGGGTAGCCATTTCACATTCTTTGAAACCAACAGGCATGAAATTCGATCGGAA	QY	1766	CGAGCTCAATATTCTACCTCTCCACACCCCCCACTATTTCCTATATACATTAATACGGT	1825
QY	686	AAAGCCTATGGCAACCGCCTAGATATTCCTCTGGCAACACGCTACGCAFTTGGGGCAGGA	Db	1561	CGAGCTCAATATTCTACCTCTCCACACCCCCCACTATTTCCTATATACATTAATACGGT	1620
Db	481	AAAGCCTATGGCAACCGCCTAGATATTCCTCTGGCAACACGCTACGCAFTTGGGGCAGGA	QY	1826	TGCAGAACACTTGAACATGCTCATGACATGCCACCATAGACAAACGCATCCCGGAGGA	1885
QY	746	CAAACCCGCAAGTGAGTTGATTCTCTTGTGGGCGNGTAAAAAAGTGATTTGGCATGAAC	Db	1621	TGCAGAACACTTGAACATGCTCATGACATGCCACCATAGACAAACGCATCCCGGAGGA	1680
Db	541	CAAACCCGTAAGTGAGTTGATTCTCTTGTGGCGGTAGTAAAAAAGTGATTTGGCATGAAC	QY	1886	TTTACAAATTTTCTCAAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA	1945
QY	806	GGCTTGTGAATAAACAATCGCGATGAACCGCATATAAACAATAAAGCGTTGACAAAGGCGAAA	Db	1681	TTTACAAATTTTCTCAAAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA	1740
Db	601	GGCTTGTGAATAAATAATTCGGGACGAAACGCGCATATAAACAAGCACTAGACAAAGGCAAAA	QY	1946	TGATATCGGTGTGATCGCATGACAAAGCTCGAATTCGCAAGCAATGGGGCGTGCAGGGA	2005
QY	866	TCTCACGGATTATCAAGTAAGAGAGACTCCCATGAAAAATGAAAAACAAGATATGTAAA	Db	1741	TGATATTCGGCGTATCGCATGACAAAGCTCGAATTCGCAAGCAATGGGGCGTGCAGGGA	1800
Db	661	TCTCACGGATTATCAAGTAAGAGAGACTCCCATGAAAAATGAAAAACAAGATATGTAAA	QY	2006	AGTGATTCCTCGAACTTTGGCAAACTGCAGAACAAAGATAAAAAAGNAATTTGGTAAGCTTC	2065
QY	926	TACCTACGGAACCAACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGA	Db	1801	AGTGATTCCTCGAACTTTGGCAAACTGCAGAACAAAGATAAAAAAGNAATTTGGTAAGCTTC	1860
Db	721	CACCTACGGAACCAACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGA	QY	2066	TGAAGATGGCAAAAGATAAACGATAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT	2125
QY	986	AGTAGAACATGACTATACCACTTATCGCGAAGAACTTAAATTTGGCGCGGTAAACTAT	Db	1861	TGAAGATGGTGAGATTAATGACAACTTCGGCATCAAAGCTATATCTCCAAATACACTAT	1920
Db	781	AGTAGAACATGACTATACCACTTATCGCGAAGAACTTAAATTTGGCGCGGTAAACTAT	QY	2126	CAACCCCGCTTTTGACCCACCGCGTGAAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	2185
QY	1046	CGGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCAC	Db	1921	TAATCCCGCTTTGACCCCATGCGGTGAGGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	1980
Db	841	CGGTGAGGGTATGGGTGAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCAC	QY	2186	CGCGCATTTGGTGTGGAAATCCTGCTTTTGGCGCTAAAAACCCAAAAATCGTATCAA	2245
QY	1106	TAAACCGATGATPATCGACTACACCGGATTTCAAAGCCGACATTTGGGATTAATAACGG	Db	1981	CGCGCATTTGGTGTGGAAATCCTGCTTTTGGTGTAAAAACCCAAAAATCGTATCAA	2040
Db	901	CAAACGATGATPATCGACTACACCGGATTTCAAAGCCGACATTTGGTATTAATAATGG	QY	2246	AGGCGGTATGGTGTCTCTCTGAAATGGGCGATTTCAACGCTCTGTGCCACCTCCCA	2305
QY	1166	CAAAATCCATGCGATTTGGCAAGCAGGAACAAGGACATGCAAGATGGCGTAAAGCCCTCA	Db	2041	AGGCGGTATGGTGTGGAAATCCTGCTTTTGGTGTAAAAACCCAAAAATCGTATCAA	2100
Db	961	CAAAATCCATGATTTTGGCAAGCGGGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA	QY	2306	ACCGGTTTATTTACCGGAAATTTTGGGCACTCAAGCGCAAGGCGAAATTTGACACCAAGCAT	2365
QY	1226	TATGGTCTGGGTGTGGGCAACAAGCACTAGCAGGGGAGGATGATTAATTAACGCTGG	Db	2101	GCCGTTTATTTACCGGAAATTTTGGGCACTCAAGGCAAGGCGAAATTTGACACCAAGCAT	2160
Db	1021	TATGGTCTGGGTGTGGGCAACAAGCACTAGCAGGGGAGGATGATTAATTAACGCTGG	QY	2366	CACTTTGTTCCTGCGCTATGAAAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCG	2425
QY	1286	GGGAATCGATTTCACACCCACTTCTCTTCTCCACAAACAATTCCTACCGCTCTAGCCAA	Db	2161	CACTTTGTTCCTGCGCTATGAAAAATGGCGTGAAGAAAAACTAGGCTTAGAGCG	2220
Db	1081	GGGGATCGATTTCACACCCACTTCTCTCTCCCAACAATTCCTACCGCTCTAGCCAA	QY	2426	CCAAGTTCTACCGGTCAAAAATCTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA	2485
QY	1346	TGCGGTTACCAACATTTGGAGGGCGGCAAGGTCCTGTAGATGGCAAGAAATGCACTAC	Db	2221	CAAGGTCTACCGGTCAAAAATCTGCGGTAAACATCACTAAGAAAGACTTCAAAATCAACAA	2280
Db	1141	TGGTGTTAACAACATTTTGGAGGTGGCAGCAGGTCCGGTAGATGGCAAGAAATGCACTAC	QY	2486	CAAAACGGCAAAATCAACCGTCGATTCGAAAAACCTTCGAGGTCTTTGTAGATGGGAAACT	2545
QY	1406	TATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTGGCGCGAGCAGAGATTTCTAT				

```
Db      2281  CRAAGCGGCGCATATCACTGTCGATCTCTAAACCTTCGAGGTCTTTGTAGATGGCAAACT 2340
Qy      2546  CTGCACTCTAAACCCACCTCGCAAGTGCCTCTAGCCACGCGTACACTTTCTTCTAGGC 2605
Db      2341  CTGCACTCTAAACCCGCTCTGAAAGTGCCTTAGCCCAAGCTACACTTTCTTCTAGGC 2400
Qy      2606  ACAAT 2610
Db      2401  ACAAT 2405

RESULT 7
LOCUS   BD185305
DEFINITION Helicobacter felis vaccine.
ACCESSION BD185305
VERSION   BD185305.1 GI:31877505
KEYWORDS JP 2002355054-A/4.
SOURCE   Helicobacter felis
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 2407)
AUTHORS  Kusters,J.G. and Cattoli,G.
TITLE    Helicobacter felis vaccine
JOURNAL  Patent: JP 2002355054-A 4 10-DEC-2002;
          AKZO NOBEL NV
COMMENT  OS Helicobacter felis
          PN JP 2002355054-A/4
          PD 10-DEC-2002
          PR 16-JUL-2001 JP 2001214711
          PR 17-JUL-2000 EP 00202565.8
          PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLI
          PC C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
          A61K39/23,
          PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
          PC C12N1/19,
          PC
          C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/ PC
          53,
          PC
          G01N33/53,G01N33/566,G01N33/569/(C12N9/80,C12R1/01),(C12Q1/68, PC
          C12R1/01),
          PC C12N15/00,C12N5/00,A61K37/02
          CC Helicobacter felis vaccine
          FH Key Location/Qualifiers
          FT CDS (2)..(682)
          FT CDS (693)..(2399).
FEATURES
source 1..2407
        /organism="Helicobacter felis"
        /mol_type="genomic DNA"
        /db_xref="taxon:214"

ORIGIN
Query Match 74.9%; Score 2158.6; DB 6; Length 2407;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy      206  GTGAACCTCACACCCAAAGCAGGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db      2  GTGAACCTCACACCCAAAGCAGGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 61
Qy      266  AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
Db      62  AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 121
Qy      326  GCCCATATTATGGACGAAGCGCGCGTGGAAAAAACCCTTGCCCAAGCTTATGGAAGAG 385
Db      122  GCCCATATTATGGACGAAGCGCGCGTGGAAAAAAGACCGTTGCGGAACCTTATGGAAGAG 181
Qy      386  TGCATGCACCTTTTGAAGAAAGATGAAGTAACTCCCGGGGTGGGTAAATATGTTGCCGAT 445

182  TGTATGCACCTTTTGAAGAAAGACGAGGTGATGCCCGGTGGGGAAATATGTCCTGAT 241
446  CTAGTGTAGAAAGCACCTTTCTCTGATGGTACGAAACTTGTAACTGTGAAATTTGGGCCATC 505
242  TTAGGCGTGGNAGTACTTTTCCGATGGCACCAAACTCGTAACGCTGAAATTTGGCCCATC 301
506  GAACCAAGATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
302  GAACCCGATGAACACTTCAAAAGCGGGCGAAGTCAAAATTTGGCTGTGATAAAGACATTGAA 361
566  CTGAATCGAGCAAAAGTAACCGAATCTTGAAGTTACTAATGAAGGGCCTAAATTCCTTG 625
362  CTCAACCGAGGTAAAGGAGTTTACCGAATAGAGTTACCAACGAAGGACCTTAAATTCCTTG 421
626  CATGTGGGTAGCCATTTCCACTTCTTGAAGCTTAACGAAGCACTAAAAATTCGATCGTGA 685
422  CATGTGGGTAGCCATTTCCACTTCTTGAAGCAACGAAGGCAATGAAATTCGATCGGAA 481
686  AAAGCCTATGCAAAACGCTAGATATTCCTCTCTGGCAACACGCTACGCAATTTGGGGCAGA 745
482  AAAGCCTATGCAAAACGCTAGATATTCCTCTCTGGCAACACGCTACGCAATTTGGGGCAGA 541
746  CAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAA 805
542  CAAACCCGTAAGTGCAGTTAATCCTCTTGGCGGCACTAAAAAGTGAATTTGGCATGAA 601
806  GGGCTTGTGAATAACATCGCGGATGAAGCCATTAACATAAAGCCCTTGACAAGCGAAA 865
602  GGGCTTGTGAATAATATTCGAGATGAAGCCATTAACATAAAGCCCTTGAAAAAGCAAAA 661
866  TCTCAGCGATTATCAAGTAAGAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA 925
662  TCTCAGCGATTATCAAGTAAGAGAGACTCCCATGAAATGAAAAACAAAGATATGTAAA 721
926  TACCTACGGACCCCAAGCGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 985
722  TACCTACGGACCTACCAAGCGCGCAAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA 781
986  AGTAGAATGATGATATACCACTATGCGGAAGAACTTAAATTTGGCGGGGTAAAACTAT 1045
782  AGTAGAATGATGATATACCACTATGCGGAAGAGCTCAAAATTTGGCGGGGTAAAACTAT 841
1046  CCGTAGGGGTATGGTCAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCATCAC 1105
842  CCGTAGGGGATGGTCAGAGCAATAGTCCAGATGAAAAACCCCTAGATTTAGTCATCAC 901
1106  TAACCGGATGATTATCGAATAACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
902  CAACCGGATGATTATGACTACACCGGATTTACAAAGCCGACATTTGCAATTTAAAAATGG 961
1166  CAAATCCATGGCATTTGGCAAGCGGAGGAAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1225
962  CAAATCCATGGCATTTGGCAAGCGGAGGAAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1021
1226  TATGGTCTGGGTGGGACAGAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
1022  TATGGTCTGGGTGGGACAGAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1081
1286  GGGATCGATTTCACACCCCACTTCTTCTTCCCAAACTTCCCTACCGCTCTAGGCCAA 1345
1082  GGGATCGATTTCACACCCCACTTCTTCTTCCCAAACTTCCCTACCGCTCTAGGCCAA 1141
1346  TGGCGTTACAACCAATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGAAATGCGACTAC 1405
1142  TGGCGTTACAACCAATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGAAATGCGACTAC 1201
1406  TATCAGCTCCGGGCAAAATGGAATCTTGCCCGCATGTTGCGCGCAGCAGAGATTTCTAT 1465
1202  CATCACTCCGGGCAAAATGGAATCTTGCCCGCATGTTGCGCGCAGCTGAAGATTTCTAT 1261
1466  GAATGTGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAAACAAGT 1525
```

Db	1262	GAATGTGGGCTTTTGGGCAAGCAATAGCTCCAGTAAAAAACAATTGTAGACAAAT	1321
Qy	1526	AGAAGCGGGCGCATTTGTTTTAAATTGTCATGAAGACTGGGGCAACAACAAGTGCAT	1585
Db	1322	AGAAGCGGGCGCATCGGCTTTAAATTGTCATGAAGACTGGGGCAACAATCCAAGTGCAT	1381
Qy	1586	CGATCACTGCTTGAGCGTGCGAGATGAATAAGATGTGCAAGTTGTGTATCACACCGATAC	1645
Db	1382	CGATCACTGCTTGAGCGTAGCAGATGAATAGATGTGCAAGTTGTATCCACACCGATAC	1441
Qy	1646	AGTCAATGAGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCA	1705
Db	1442	GGTCAATGAGGCAAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCA	1501
Qy	1706	TGCTACCAATGAGGAGCGGGTGAGGACACTCACTGATGTTATCACCATGGCAGG	1765
Db	1502	TGCTACCAATGAGGAGCGGGTGAGGACACTCACTGATGTTATCACCATGGCAGG	1561
Qy	1766	CGAGCTCAATATTCTACCTTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT	1825
Db	1562	CGAGCTCAATATTCTACCTTCTCCACCAACCCCACTATCCCTATACCAATTAATACGGT	1621
Qy	1826	TGCAGAACACTTAGACATGCTATGATGCCACCACTAGACAAACGCAATCCGCGAGGA	1885
Db	1622	TGCAGAACACTTAGACATGCTATGATGCCACCACTAGACAAACGCAATCCGCGAGGA	1681
Qy	1886	TTTACAATTTTCTCAAAAGCCGATTCGCCGCCGGCTCTATCGGGCTCAAGATGTGCTCCA	1945
Db	1682	TTTACAATTTTCCAAAGCGTATCCGCCGCCGGCTCTATCGCGCTGAAAGATGTGCTCCA	1741
Qy	1946	TGATATCGGTGTGATCCCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCGAGCGA	2005
Db	1742	TGATATTTGGGTGATTCGCGATGACAAGCTCGGATTCGCAAGCAATGGGCGTGCGAGCGA	1801
Qy	2006	AGTGATTTCTCGAACTTTGGCAGACTGCGGATAGAAATAAAAAGAAATTTGGTAAAGCTTCC	2065
Db	1802	AGTGATTTCTCGAACTTTGGCAGACTGCGGATAGAAATAAAAAGAAATTTGGTAAAGCTTCC	1861
Qy	2066	TGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGGTACATCTCCAAATACACTAT	2125
Db	1862	TGAAGATGTCAGATAACGACAATTTCCGATTCAAAGCGTACATCTCCAAATACACTAT	1921
Qy	2126	CAACCCCGCTTTGACCCACGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT	2185
Db	1922	TAACCCCGCTCTAACCCCATGSGGTAAAGCAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT	1981
Qy	2186	CGCCGACTTGGTGTGTGGAATCTCGCTTTTTTTTGGCGTAAACCCCAAAATCTGTGATCAA	2245
Db	1982	CGCTGTATTTGGTGTGTGGAATCTCGCTTTTTTTTGGTGTGAAACCTAAGATTTGTGATCAA	2041
Qy	2246	AGGCGGTATGGTGTCTTCTGTAATGGGCGATTTCTAAACGCGTCTGTGCCACTCCCCCA	2305
Db	2042	AGGCGGTATGGTGTCTTCTGTAATGGGCGACTTCAAACGCGTCCGTCCTACACCTCA	2101
Qy	2306	ACCGGTTTTATTACCGCGAAATGTTTTGGGCATCACGGCAAGCGCAAAATTTGACACCAAGAT	2365
Db	2102	GCCGGTTTTATTACCGCGAAATGTTTTGGGCATCACGGCAAGCGCAAAATTTGACACCAAGAT	2161
Qy	2366	CACTTTTGTTTTCAAAGTTCGCTATGAAATTTGGCGTGAAGAAAGCTGGGCTTAGAGCG	2425
Db	2162	CACTTTTGTTTTCAAAGTTCGCTATGAAATTTGGCGTGAAGAAAGCTTAGAGCG	2221
Qy	2426	CCAAAGTTCTACCGGTCAAAAACCTGCGGTAAACATCACCAAGAAAGACATTCAGTTTCAACGA	2485
Db	2222	CAAGGTGCTACCCGTGAAAAACCTGCCGCAACATCACTAAGAAAGACATTCAAATTCACAA	2281
Qy	2486	CAAAAACGCAAAAAATCACCGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACT	2545
Db	2282	CAAGACGGCGCATATCACTGTCGATCTTAACACCTTCGAGGTCTTTGTAGATGGCAAACT	2341
Qy	2546	CTGCACTCTTAAACCCACCTCGCAAGTGCCTCTTAGCCAGCGCTACACTTTCTTTAGGC	2605
Db	2342	CTGCACTCTTAAACCCGCTCTGAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTTAGGC	2401

[illegible]

QY 386 TGCATGCACCTTTTGTAAAAAGATGAAGTAATGCCCCGGGTGGTAATATATGTTCCCGAT 445
DB 182 TGTATGCACCTTTTGTAAAAAGAGCAGGTGATGCCCGTGTGGGAAATATGTTCCCTCAT 241
QY 446 CTAGGTGTAGAAGCCACTTTCCTGATGAGTAACGAACTTGTAACTGTGAATTTGGCCCATC 505
DB 242 TTAGGCGTGAAGCTACTTTCCGATGGCACCAACTCGTAACCGTGAATTTGGCCCATC 301
QY 506 GAAACCCAGATGAGCACTTCAAAGCGGGCGAAAGTGAATTTGGTTCCGATAAAGACATCGAG 565
DB 302 GAACCCGATGAACTTCAAAGCGGGCGAAAGTCAAATTTGGTGTGATTAAGACATTTGAA 361
QY 566 CTCAATGCAGCGCAAGAAAGTAACGAACTTTGAAGTTTACTAATGAAGGCCCTAAATCCTTG 625
DB 362 CTCAAACGAGTGAAGGAATTAACGAACTAGAAGTTACCAACGAAGCACTTAAATCCTTG 421
QY 626 CATGTGGGTAGCCATTTCACCTTCTTTGAAGCTTAAAGGCACTTAAATTCGATCGTGAA 685
DB 422 CATGTGGGTAGCCATTTCACCTTCTTTGAAGCCAAACGAAGGCATTTGAATTCGGGAA 481
QY 686 AAAGCCTATGCGAAACCGCTAGATATTCCCTCTGGCAACAGCTACGCATTGGGGCAGGA 745
DB 482 AAAGCCTATGCGAAACCGCTAGATATTCCCTCTGGCAACAGCTACGCATTGGGGCAGGA 541
QY 746 CAAACCCGCAAGTGCAAGTTGATTCCTCTTGGTGCGCAAGTAAAGAGTGAATGGCATGAAC 805
DB 542 CAAACCCGTAAGTGCAAGTTAATCCCTCTTGGCGCAGTAAAGAGTGAATGGCATGAAC 601
QY 806 GGCCTTGTGAATTAACATCGCGGATGAACGCGCATAAACATTAAGCGCTTGACAAGCGCAAA 865
DB 602 GGCCTTGTGAATTAATTTGCAGATGAACGCGCATAAACATTAAGCGCTTGAGAAAAGCAAAA 661
QY 866 TCTCAGGATTTATCAAGTAAGAGACTCCCATGAAATGAAGAAACCAAGAAATGTAATA 925
DB 662 TCTCAGGATTTATCAAAATAGGAGACTCCCATGAAATGAAGAAACCAAGAGTATGAAA 721
QY 926 TACCTACGGACCCACCAAGCGCATAAAGTCGCTTAGGAGATACCCGATCTTTGGGCAGA 985
DB 722 TACCTACGGACCTTACCACGCGCAAAAGTGGCTTAGGAGATACCCGATCTTTGGGCAGA 781
QY 986 AGTAGAACATGATATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTAAAACTAT 1045
DB 782 AGTAGAACATGATATACCACTTATGGCGAAGAGCTCAAATTTGGCGCGGTAAAACTAT 841
QY 1046 CCGTAGGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACCCCTAGATTTAGTCATCAC 1105
DB 842 CCGTAGGGATGGGTGAGAGCAATAGTCAGATGAAAAACCCCTAGATTTAGTCATCAC 901
QY 1106 TAAACGCGATGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
DB 902 AAACGCGATGATTATTTGACTACACCGGATTTTAAAGCCGACATTTGGCAATTTAAAAATGG 961
QY 1166 CAAAATCCATGGCAITTTGGCAAGGCAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1225
DB 962 CAAAATCCATGGCAITTTGGCAAGGCAGGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCA 1021
QY 1226 TATGTCGTGGGTGAGGCAAGAGCACTAGCAGGGGAAAGGTATGATTAATTCAGCTGG 1285
DB 1022 TATGTCGTGGGTGAGGCAAGAGCACTAGCAGGGGAAAGGTATGATTAATTCAGCTGG 1081
QY 1286 GGGAAATCGAATTCACACCCACTTCCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1345
DB 1082 GGGAAATCGAATTCACACCCACTTCCCTTCTCCACAACTTCCCTACCGCTCTAGCCAA 1141
QY 1346 TGGCGTTACAAACCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCAAGATGGCACTAC 1405
DB 1142 TGGCGTTACAAACCATGTTTGGCGGTGGCAGAGTCCGTTAGATGGCAAGATGGCACTAC 1201
QY 1406 TATCACTCCGGGCAAAATGGAACCTTGACCGCATTTGTCGCGCAGCAGAGAGTATTCAT 1465
DB 1202 CATCACTCCGGGCAAAATGGAACCTTGACCGCATTTGTCGCGCAGCTGAAGAGTATTCAT 1261
QY 1466 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACCACTTTGTAGAACAGT 1525

DB 1262 GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCCAGTAAAAAACCACTTTGTAGAACAAAT 1321
QY 1526 AGAAGCGGGCCGATTTGGTTTAAATTCATGAAGACTGGGGCAAAACCAAGTGGCAT 1585
DB 1322 AGAAGCGGGCCGATCGGCTTTAAATTCATGAAGACTGGGGCAAACTTCAAGTGCAAT 1381
QY 1586 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGCAAGTTTGTATCCACACCGATAC 1645
DB 1382 CGATCACTGCTTGAGCGTGGCAGATGAATACGATGCAAGTTTGTATCCACACCGATAC 1441
QY 1646 AGTCAATGAGCAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGGCCATCCA 1705
DB 1442 GGTCAATGAGCAGGTTATGTAGATGACACCTTGAATGCAATGAACGGCGGCCATCCA 1501
QY 1706 TGCCTACCACTTGAAGGAGCGGTGGAGGACACTCACTGATGTTTATCACCATTGGCAGG 1765
DB 1502 TGCCTACCACTTGAAGGAGCGGTGGAGGACACTCACTGATGTTTATCACCATTGGCAGG 1561
QY 1766 CGAGCTCAATTTCTACCTCTCCACCAACCCCACTATTCCTATACCAATTAACGGT 1825
DB 1562 CGAGCTCAATTTCTACCTCTCCACCAACCCCACTATTCCTATACCAATTAACGGT 1621
QY 1826 TGCAGAACACTTGAAGCATGCTCATGACATGCCACCACTAGACAAACGGCATCCGCGAGA 1885
DB 1622 TGCAGAACACTTGAAGCATGCTCATGACATGCCACCACTAGATAAACGCAATCCGCGAGA 1681
QY 1886 TTTTCAATTTTCTCAAGCGGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA 1945
DB 1682 TTTTCAATTTTCCCAAGCCGTATCCGCCCGGCTCTATCCGGCTGAAGATGTGCTCCA 1741
QY 1946 TGATATGGGTGATCGCATGAAGAAGCTCCGATTCGCAAGCAATGGGGCGTGCAGGCGA 2005
DB 1742 TGATATTTGGCGTGAATCGCATGACAGCTCGATTCGCAAGCAATGGGGCGTGCAGG 1801
QY 2006 AGTGATTTCTCGAATTTGGCAGACTGCGGATAAGAAATAAAAAGAAATTTGTAAGTTC 2065
DB 1802 AGTGATTTCTCGAATTTGGCAGACTGCGGATAAGAAATAAAAAGAAATTTGTAAGTTC 1861
QY 2066 TGAAGATGGCAAGATAACGATAATTTCCGATTAAGCGCTACATCTCCAATATACATAT 2125
DB 1862 TGAAGATAGTGACATAACGACAACTTCCGTATCAACGCTACATCTCCAATATACATAT 1921
QY 2126 CAACCCCGCTTTGACCCACGCGTGAAGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
DB 1922 TAACCCCGCTCTAACCCATGGGGTAAGCGATATATCGGCTCTGTGGAAGAGGCGCAAT 1981
QY 2186 CGCCGACTTGTGTGTGGAAATCCTGCTTTTGGCGTAAACCCCAAAATCGTGATCAA 2245
DB 1982 CGCTGATTTGGTGTGTGGAAATCCTGCTTTTGGTGTGAAACCTAAGATTTGTATCAA 2041
QY 2246 AGGCGGTATGGTGTCTTCTGTAAATGGGCGATTTCTAAAGCGGTCTGTGCCACCTCCCA 2305
DB 2042 AGGCGGTATGGTGTCTTCTGTAAATGGGCGATTTCTCAACGCGTCCGTGCTTACCTCA 2101
QY 2306 ACCGTTTATACCGCGAAATTTTGGGCATCAAGCAAGCGGAAATTTGACACCGCAT 2365
DB 2102 GCGGTTTATACCGCGAAATTTTGGGCATCAAGCAAGCGGAAATTTGACACCGCAT 2161
QY 2366 CACTTTTGTTCCAAAGTCGCTTATGAAAATGGGCTGAAAGAAAAGCTGGGCTTAGAGCG 2425
DB 2162 CACTTTTGTTCCAAAGTCGCTTATGAAAATGGGCTGAAAGAAAAGCTTAGAGCGG 2221
QY 2426 CCAAGTTCTACCGGTCAAAAATCGCGGTAAACATCAACCAAGAAAGACTTTCAAGTTCAA 2485
DB 2222 CAAGGTGTACCCCGTGAAAAATCTCGCGCAACATCACTAAGAAAGACTTTCAAATTTCAA 2281
QY 2486 CAAAACGGCAAAAATCAACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACT 2545
DB 2282 CAAGACGGCGCATATCACTGTCTTCTTAAACCTTCGAGGTCTTTGTAGATGGCAACT 2341
QY 2546 CTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTCTTCTTAGGC 2605

Db	2342	CTGCACCTCTAAACCGCGCTCTGAAGTGCTCTAGCCAGCGCTACACTTTCTTCTTAGCG	2401
Qy	2606	ACAATG 2611	
Db	2402	NCAATG 2407	
RESULT 9			
BD185304			
LOCUS		2183 bp	DNA
DEFINITION	BD185304	Helicobacter felis vaccine.	
ACCESSION	BD185304		
VERSION	BD185304.1	GI:31877504	
KEYWORDS	JP 2002355054-A/3.		
SOURCE	Helicobacter felis		
ORGANISM	Helicobacter felis		
REFERENCE		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.	
AUTHORS		1 (bases 1 to 2183)	
TITLE		Kusters, J.G. and Cattolli, G.	
JOURNAL		Helicobacter felis vaccine	
		Patent: JP 2002355054-A 3 10-DEC-2002;	
		AKZO NOBEL NV	
COMMENT		OS Helicobacter felis	
		PN JP 2002355054-A/3	
		PD 10-DEC-2002	
		PF 16-JUL-2001 JP 2001214711	
		PR 17-JUL-2000 EP 00202565.8	
		PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOLI	
		PC	
		C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC	
		A61K39/23,	
		PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,	
		PC C12N1/19,	
		PC	
		C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC	
		53,	
		PC	
		G01N33/53 G01N33/566, G01N33/569// (C12N9/80, C12R1:01), (C12Q1/68, PC	
		C12R1:01),	
		PC C12N15/00, C12N5/00, A61K37/02	
		CC Helicobacter felis vaccine	
		FH Key	
		Location/Qualifiers	
		FT CDS (3)..(683)	
		FT CDS (694)..(2181).	
FEATURES		Location/Qualifiers	
source		1..2183	
		/organism="Helicobacter felis"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:214"	
ORIGIN			
Query Match		67.6%; Score 1948; DB 6; Length 2183;	
Best Local Similarity		93.3%; Pred. No. 0;	
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;			
Qy	206	GTGAACCTCACACCCAAAGAGCGGCTTAAAGCTCAACCAACCGGCAATTCCTTACATTAGT	325
Db	3	GTGAACCTCACACCCAAAGAGCGGCTTAAAGCTCAATCAACCGGCAATTCCTTACATTAGT	62
Qy	266	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGGCAATTCCTTACATTAGT	325
Db	63	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCGGCAATTCCTTACATTAGT	122
Qy	326	GCCCATATTATGGAAGAGCGGCGCGTGGAAAAAACCCTTCCAGCTTATGGAAGAG	385
Db	123	GCCCATATTATGGAAGAGCGGCGCGTGGAAAAAACCCTTCCAGCTTATGGAAGAA	192
Qy	386	TGATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGGTGGGTAAATATGGTTCCCGAT	445
Db	183	TGATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGGTGGGTAAATATGGTTCCCGAT	242
Qy	446	CTAGGTGTAGAGCCACCTTCTCTGATGGTACGAACCTTGTAACCTGTGAATTGGCCCATC	505
243	TTGGCGGTAGAACCACTTTCCCGGATGGCAACCAACTCGTAACGTAATTTGCCCAATT	302	
506	GAACACAGATGAGCACTTCAAAAGCGGCGAAGTGAATTTGGTTCGATTAAGACATCCAG	565	
303	GAACCTGATGAACACCTTTAAAGCGGCGTGAAGTGAATTTGGCTGTATTAAGACATTTGAG	362	
566	CTCAATGCGAGCAAAAGATTAACCGAATTTGAGGTTACTAATGAAGGCGCTTAAATTCCTTG	625	
363	CTCAACGTGGGTAAGGAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTTAAATTCCTTG	422	
626	CATGTGGGTAGCCATTTTCCACTCTTTTGAAGCTAAACAAGGCACTAAATTCGATCGTGAA	685	
423	CATGTGGGTAGCCATTTTCCACTCTTTTGAACCAACCAAGGCACTTGAATTCGATCGGGAA	482	
686	AAAGCCTATGGCAAAACGCTTAGATATTCCTCTCTGGCAACGCTACCGATTTGGGGCAGGA	745	
483	AAAGCCTATGGCAAAACGCTTAGATATTCCTCTCTGGCAACGCTACCGATTTGGGGCAGGA	542	
746	CAAAACCGCAAAAGTGCAGTTGATTTCCTCTTGGTGGCAGTAAAAAGTGAATTTGGCATGAAC	805	
543	CAAAACCGTAAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAGTGAATTTGGCATGAAC	602	
806	GGGCTTGTGAATTAACATCGCGGATGAACGCGCATAAACATAAAGCGCTTGAACAAGCGGAAA	865	
603	GGGCTTGTGAATTAATATTCGGGACGAACGCGCATAAACACAAAGCACTAGACAAGGCAAAA	662	
866	TCTCAGGATTTATCAAGTAAGGAGCTCCCATGAAATGAATAAACAAGTAATGTATAA	925	
663	TCTCAGGATTTATCAAGTAAGGAGCTCCCATGAAATGAATAAACAAGTAATGTATAA	722	
926	TACTACGGACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA	985	
723	CACCTACGGACCCACCAAGGCGATAAAGTGCCTTAGGAGATACCGATCTTTTGGGCAGA	782	
986	AGTAGAACATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGGGGTAAAAACTAT	1045	
783	AGTAGAACATGACTATACCACTTATGGCGAAGAGCTCAATTTTGGCGGGGTAAAAACTAT	842	
1046	CCGTAGGGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACACCTAGATTTAGTCATCAC	1105	
843	CCGTAGGGGTATGGGTGAGAGCAATAGCCCTGTATGAAAAACACCTTAGATTTAGTCATCAC	902	
1106	TAACGCGATGATTATCAGCTACACCGGATTTACAAAGCGGACATTTGGGATTTAAAAACGG	1165	
903	CAACGCGATGATTATCAGCTACACCGGATTTTAAAGCCGACATTTGGTATTTAAAAATGG	962	
1166	CAAAATCCATGGCATTTGGCAAGGCGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCA	1225	
963	CAAAATCCATGGCATTTGGCAAGGCGGGAACAAGACATGCAAGATGGCGTAAGCCCTCA	1022	
1226	TATGTGCTGGGTGTGGGCAAGGACATGACAGGGGAAGGTATGATTTATCCGCTGG	1285	
1023	TATGTGCTGGGTGTGGGCAAGGACATGACAGGGGAAGGTATGATTTATCCGCTGG	1082	
1286	GGGAATCGATTTCACACCCACTTCCCTTTCTCCACAACAAATTCCTCCTAGCCAA	1345	
1083	GGGAATCGATTTCACACCCACTTCCCTTTCTCCACAACAAATTCCTCCTAGCCAA	1142	
1346	TGGCGTTTACAAACCATTTTGGAGCGGCGACAGGTCCTGTAGATGGCAGAAATGCCACTAC	1405	
1143	TGGTGTACAAACCATTTTGGAGGTGGCAAGTCCGCTAGATGGCAGAAATGCCACCAC	1202	
1406	TATCAGCTCCGGGCAAAATGGAACTTTGCACCGCATGTTGGCGCAGCAGAGAGTATTTCTAT	1465	
1203	CATCAGCTCCGGGCAAAATGGAACTTTGCACCGCATGTTGGCGCAGCTGAAGAGTATTTCTAT	1262	
1466	GAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT	1525	
1263	GAATGTAGGGCTTTTGGGCAAAAGGCAATAGTCTTAGCAAAAAACAATTTGTAGAACAGT	1322	
1526	AGAAGCGGCGCATTTGGTGTAAATTTAAATTTGCAATGAAGACTGGGGCAACAACAAGTGGCAT	1585	

Db	1323	AGAAGCGGGCGCGATTGGCTTTAAATTGCAATGGAAGACTGGGGCACAAACACAAAGTCGCAT	1382
Qy	1586	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1645
Db	1383	CGATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1442
Qy	1646	AGTCAATAGCGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCGCATCCA	1705
Db	1443	GGTCAATAGCGCAGGTTATGTGATGACACCTTAATGCAATGAACGGGCGCGCATCCA	1502
Qy	1706	TGCTTACCACATTGAGGAGCGGGTGGAGGACATCACTGATGTTATCAACCATGGCAGG	1765
Db	1503	TGCTTACCACATTGAGGAGCGGGGAGGACATCACTGATGTTATCAACCATGGCAGG	1562
Qy	1766	CGAGCTCAATATTCTACCTCTCCACACCCACATATTCCCTATACCAATTAATACGGT	1825
Db	1563	CGAGCTCAATATTCTACCTCTCCACACCCACATATTCCCTATACCAATTAATACGGT	1622
Qy	1826	TGCAGAACACTTAGACATGCTCATGACATGCCACCTAGACAAACGCATCCGCGAGGA	1885
Db	1623	TGCAGAACACTTAGACATGCTCATGACCTGCCACCTAGATAAGCGCATCCGCGAGGA	1682
Qy	1886	TTTACAAATTTTCTCAAGCCGATATCCGCCCGGCTCTATCGCGCTGAAGATGTCTCCA	1945
Db	1683	TTTACAAATTTTCTCAAGCCGATATCCGCCCGGATCTATTGCCGCTGAGGATGTCTCCA	1742
Qy	1946	TGATATGGGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGGCTGCAGCGGA	2005
Db	1743	TGATATGGGTGATCGCGATGACTAGCTCCGATTCGCAAGCAATGGGGCTGCAGCGGA	1802
Qy	2006	AGTGATTCCTCGAACTTGGCAGACTCGGATAAGATAAAGAAATTTGGTAAAGCTTCC	2065
Db	1803	AGTGATTCCTAGAACTTGGCAACTGCAGATAGATATAAAGAAATTTGGTAAAGCTTCC	1862
Qy	2066	TGAAGATGGCAAGATAACGATAATTTCCGCTATTAAGCGCTACATCTCCAATACACTAT	2125
Db	1863	TGAAGATGGTGAGATAACGACAACTTCCGCTATTAAGCGCTATATCTCCAATACACTAT	1922
Qy	2126	CAACCCCGCTTACCCCAACCGGCTGAGCGATATATCGGCTCTGTGGAAGAGGCGAAGAT	2185
Db	1923	TAATCCCGCTTACCCCAACCGGCTGAGCGATATATCGGCTCTGTGGAAGAGGCGAAGAT	1982
Qy	2186	CGCCGACTTGGTGGTGAATCTCGCTTTTGGGTAAAGCCCAAAATCTGTATCAA	2245
Db	1983	CGCCGACTTGGTGGTGAATCTCGCTTTTGGGTGAACCTTAAGATTTGTATTA	2042
Qy	2246	AGCGGTATGGTGGTCTTCTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCCCCA	2305
Db	2043	AGGTGGCATGGTGGTCTTCTCTGAAATGGCGGATTTCTAACCGCTCTGTGCCCACTCCA	2102
Qy	2306	ACCGGTTTATACCGGAAATGTTTGGGATCAACGGCAAGCGGAAATTTGACACCGAGCAT	2365
Db	2103	GCCGGTTTATACCGGAAATGTTTGGGATCAACGGCAAGCGGAAATTTGACACCGAGCAT	2162
Qy	2366	CACTTTGTGTTCCAAAGTCG	2385
Db	2163	CACTTTGTGTTCTCAAGCG	2182

RESULT 10
AX356689
LOCUS AX356689
DEFINITION Sequence 7 from Patent EP1176192.
ACCESSION AX356689
VERSION AX356689.1 GI:18674026
KEYWORDS Helicobacter felis
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1
AUTHORS Kusters, J.G. and Cattoli, G.
TITLE Helicobacter felis vaccine

JOURNAL	Patent: EP 1176192-A 7 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1. .2183
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	3. .683
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23259.1"
	/db_xref="GI:18674027"
CDS	694. .>2181
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23260.1"
	/db_xref="GI:18674028"
	/translation="MKKKQKQYVNTYPTTGGDKVRLGDTDLWAEVHDYTTTGYBELKFGAGKTIREGMGQSPDENTLVLITNAMIIDYTIYKADIGIKNGKIHGKAGKAGNDMDGVSPHVMVVGQTEALAGEBMIITAGGIDSHFSLSPQOFTKALANGVITMPGGGTGPDVGTNATITPGKWNHMLRAAEYSMNKVPFLGKGNSSSKQOLVQVQVGAAGIHFGLHEDWGTPPSAIDHCLSVADYDQVCIHTPTINTVAEHLMLTCHLDKRIREDIQFSAGGSPDVTIMAGELNLPSTPTTPTTINTVAEHLMLTCHLDKRIREDIQFSQSRIRPGSIAEEDVLHDIGVIATMSDSQAMGRAGEVIRPTWQADKNKGFGLKPEDGANDNPRIKRYISKYITNPALTHGVSEYIGSVEEGIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGSDNSAVPTPQPVVYREMFHGHGKAKFDTSITFRVSS"
ORIGIN	
Query Match	67.6%; Score 1948; DB 6; Length 2183;
Best Local Similarity	93.3%; Pred. No. 0;
Matches 2035; Conservative	0; Mismatches 145; Indels 0; Gaps 0;
Qy	206 GTGAACTCAACCCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 265
Db	3 GTGAACTCAACCCAAAGAGCAAGAAAGTTCTTGTATATATATGCGGGCGAAGTGGCT 62
Qy	266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAAGT 325
Db	63 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCATTGCTTACATTAAGT 122
Qy	326 GCCCATATTATGGAACGAGCGCGCTGGAAAAAACCCTTGGCCAGCTTATGGAAGAG 385
Db	123 GCCCATATTATGGAACGAGCGCGCTGGAAAAAACCCTTGGCTGAACCTTATGGAAGAA 182
Qy	386 TGCATGACATTTTGAAGAAAGATGAAGTAATGCCCGGGTGGGTAAATATGGTTCCCGAT 445
Db	183 TGTATGACATTTTGAAGAAAGATGAGTGTATGCCCGGTGGGGAATATGGTCCCTGAT 242
Qy	446 CTAGGTGTAGACCCACTTTCTGATGTTAGCAAACTTGTAACTGTGAATTTGGCCCATC 505
Db	243 TTGGGCGTAGAGCCACTTTCCCGGATGGCAACCAACTCGTAACCGTGAATTTGGCCCAT 302
Qy	506 GAACCATGAGCACTTCAAGCGGGCGAAGTGAATTTTGGTTGCGATAAAGACATCGAG 565
Db	303 GAACCTGATGAACACTTTAAAGCGCGTGAAGTGAATTTTGGCTGTGATAAAGACATTTGAG 362
Qy	566 CTCAATGAGCGCAAGAAAGTAACCGAACTTGAGGTGTACTAATGAAGGGCCTAAATCCTTG 625
Db	363 CTCAACGTGGTAAAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATCCTTG 422
Qy	626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTTAACAGGCACTAAATTTGATGTCGTGAA 685
Db	423 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTTAACCAACGAAGGCACTTGAATTCGATCGGAA 482
Qy	686 AAAGCCTATGGCAACCGCTAGATATTCCCTCTGGCAACACGCTACGCAATTTGGGCGAGGA 745

Db	483	AAAGCCTATGGCCAAAGCCCTAGATATTCCTCTGGCAACCGCTACGCATTGGGGCAGGA	542
QY	746	CAAAACCCGCAAAAGTCAGTTGATTCTCTTGGTGGCAGTAAAAAAGTGATTTGGCATGAAC	805
Db	543	CAAAACCCGTAAGTGACGTTAATCCCTCTTGGCGGTAGTAAAAAAGTGATTTGGCATGAAC	602
QY	806	GGGCTTGTGAATAACATCCGGATGAACGCCATAAATCATTAAGCGCTTGACAAGGGCAAA	865
Db	603	GGGCTTGTGAATAATATTGCGACGAAACGCCATAAAACAAAAGCACTAGACAGGCAAAA	662
QY	866	TCTCAGCGATTATCAAGTAAGGAGACTCCCATGAAATGAAATAACAAAACAGAATATGTAAA	925
Db	663	TCTCAGCGATTATCAAGTAAGGAGACTCCCATGAAATGAAATAACAAAACAGAATATGTAAA	722
QY	926	TACCTACGGACCCCAAAAGGGCGATTAAGTGGCTTTAGGAGATACCGATCTTTTGGGCAGA	985
Db	723	CACCTACGGACCCCAACACAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCAGA	782
QY	986	AGTAGAACATGACTATACCACTATGGCGAAGAACTTTAAATTTGGCGCGGGTAAAACTAT	1045
Db	783	AGTAGAACATGACTATACCACTATGGCGAAGAGCTCAAAATTTGGCGCGGGTAAAACTAT	842
QY	1046	CGGTGAGGGTATGGGTACAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATCAC	1105
Db	843	CGGTGAGGGTATGGGTACAGCAATAGCCCAAGATGAAACACCCCTAGATTTAGTCATCAC	902
QY	1106	TAAACGCGATGATATTCGACTACACCCGGGATTTACAAAGCCGACATTTGGGATTAATAACGG	1165
Db	903	CAACGCGATGATATTCGACTACACCCGGGATTTATAAGCCGACATTTGGTATTAATAATGG	962
QY	1166	CAAAATCCATGGCATTTGGCAAGCAGGAAACAAAGCATGCAAGATGGCGTAAAGCCCTCA	1225
Db	963	CAAAATCCATGGTATTGGCAAGCAGGAAACAAAGCATGCAAGATGGCGTAAAGCCCTCA	1022
QY	1226	TATGTCGTGGGTGGGCACAGAACACTAGCAGGGAGGTATGATTTATTCACCGCTGG	1285
Db	1023	TATGTCGTGGGTGGGCACAGAACACTAGCAGGGAGGTATGATTTATTCACCGCTGG	1082
QY	1286	GGGAATCCGATTCACACACCCACTTCCTTCTCCACAAACAAATCCCTACCGCTCTAGCCAA	1345
Db	1083	GGGGATCGATTCCACACCCACTTCCTTCTCCCAACAAATCCCTACCGCTCTAGCCAA	1142
QY	1346	TGGCGTTACAAACCATGTTTGGAGCGGCACAGGTCTCTGTAGATGGCACGAAATCGACTAC	1405
Db	1143	TGGTGTACAAACCATGTTTGGAGTGGCACAGGTCCGTTAGATGGCACGAAATCGACTAC	1202
QY	1406	TATCACTCCGGCAATAGAACTTGGCACCGCATGTTGGCGGCAGCAGAGAGATTCTAT	1465
Db	1203	CATCACTCCGGCAATAGAACTTGGCACCGCATGTTGGCGGCAGCTCAAGAGATTCTAT	1262
QY	1466	GAATGTGGCTTTTGGGCAAAAGGCAATAGCTCTAGCNAACAACTTGTAGAACAGT	1525
Db	1263	GAATGTAGGCTTTTGGGCAAAAGGCAATAGTCTAGCNAACAACTTGTAGAACAGT	1322
QY	1526	AGAAGCGGGCGGATTTGGTTTAAATGCAATGAAGACTGGGGGCACACACCAAGTGGCAT	1585
Db	1323	AGBAGCGGGCGGATTTGGCTTTAAATGCAATGAAGACTGGGGGCACACACCAAGTGGCAT	1382
QY	1586	CGATCACTGCTTGGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1645
Db	1383	CGATCACTGCTTGGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1442
QY	1646	AGTCAATGAGCAGGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCA	1705
Db	1443	GGTCAATGAGCAGGTATGTAGATGACACCCCTAAATGCAATGAACGGGCGGCCATCCA	1502
QY	1706	TGCTTACCACATTGAGGGCGGGTGGAGGACACTCACCTGATGTTATACCATGGCAGG	1765
Db	1503	TGCTTACCACATTGAGGGCGGGCGGAGGACACTCACCTGATGTTATACCATGGCAGG	1562
QY	1766	CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT	1825
Db	1563	CGAGCTCAATATTCTACCTCTCCACCAACCCCACTATTCCCTATACCAATTAATACGGT	1622

QY	1826	TGAGAAACACTTAGACATGCTATGACATGCCACCACCTAGACAAACGCAATCCGCGAGGA	1885
Db	1623	TGAGAAACACTTAGACATGCTATGACATGCCACCACCTAGATAAGCGCATCCGCGAGGA	1682
QY	1886	TTTACAAATTTCTCAAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAGATGCTCTCA	1945
Db	1683	TTTACAAATTTCTCAAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAGATGCTCTCA	1742
QY	1946	TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGA	2005
Db	1743	TGATATTTGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCGA	1802
QY	2006	AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCC	2065
Db	1803	AGTGATTTCTCGAACTTTGGCAGAACTGCGAGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCC	1862
QY	2066	TGAAGATGGCAAGATTAACGATTAATTTCCGCAATTAAGCGGTACATCTCCAATACACTAT	2125
Db	1863	TGAAGATGGTGCAGATAACGCAAACTTTCCGCAATTAAGCGGTACATCTCCAATACACTAT	1922
QY	2126	CAACCCCGCTTTGACCCACCGCGTACGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	2185
Db	1923	TAATCCCGCTTTGACCCATGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGAT	1982
QY	2186	CGCGGACTTGGTGGTGGGAATCCTGCTTTTGGCGTAAAAACCCAAATCGTGATCAA	2245
Db	1983	CGCGGACTTGGTGGTGGGAATCCTGCTTTTGGCGTAAAAACCCAAATCGTGATCAA	2042
QY	2246	AGCGGTATGGTGTCTTCTCGAAATGGGGATTTCAAACGCTCTGTGCCACTCCCA	2305
Db	2043	AGTGCGCATGGTGTCTTCTCGAAATGGGCGATTTAAACGCGTCCGTCGCCACGCTCA	2102
QY	2306	ACCGGTTTATTACCGCGAAATGTTGGGCGATCACGCAAGCGAAATTTGACACCGCAT	2365
Db	2103	GCCGTTTATTACCGCGAAATGTTGGGCGACCGCAAGCGAAATTTGACACCGCAT	2162
QY	2366	CACCTTTGTTTCCAAAGTCG	2385
Db	2163	CACCTTTGTTCTCAAGCG	2182

RESULT 11	AF330621	8406 bp	DNA	linear	BCT 13-MAY-2003
LOCUS	AF330621				
DEFINITION	Helicobacter bizzozeronii tRNA ribosyltransferase gene, partial cds; GIB, complete cds; urease gene cluster, complete sequence and unknown gene.				
ACCESSION	AF330621				
VERSION	AF330621.1	GI:27462193			
KEYWORDS	Helicobacter bizzozeronii				
SOURCE	Helicobacter bizzozeronii				
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.				
REFERENCE	1 (bases 1 to 8406)				
AUTHORS	Zhu,J., Teng,C.H., Chang,C.F., Chang,C.D., Simpson,K.W., Wei,C., McDonough,P., McDonough,S. and Chang,Y.F.				
TITLE	Cloning and characterization of a Helicobacter bizzozeronii urease gene cluster				
JOURNAL	DNA Seq. 13 (6), 321-331 (2002)				
PUBMED	12652903				
REFERENCE	2 (bases 1 to 8406)				
AUTHORS	Zhu,J. and Chang,Y.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-DEC-2000) Population Medicine and Diagnostic Science, College of Veterinary Medicine, Cornell University, Tower Road, Ithaca, NY 14853, USA				
FEATURES	Location/Qualifiers				
source	1. .8406				
	/organism="Helicobacter bizzozeronii"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:56877"				

QY 493 GAAATGGCCCATCGAACCAAGATGAGCACTTTCAAAGCGGGGGAAGTGAAATTTGGTGGCGA 552
Db 2628 CCATATCCCCCGTTGAAGATGAGTGGGCAATAAATTGGCTCCGGGTGAAGTGATTTTGA AAAA 2687
QY 553 TAAAGACATCGAGCTCAATCGAGGCAAGAGTAACCGAACTTTGAGGTTACTTAATCAAGG 612
Db 2688 CGAAGACATCACTTTGAATTCGAGGCAAAACAAGCCACCTTTAAGAAGTGATTAACAAGG 2747
QY 613 GCCTAAATCTTTGCATGTGGTAGCCATTTCCACTTTCTTTGAAGCTAAACAAGCACTTAA 672
Db 2748 CGATCGCCCGGTGCAAGTGGCTCCCACTTCCACTTTCTTTGAAGTGAATAAGCTTTTGA 2807
QY 673 ATTGCATGCTGAAAAAGCCTATGCGAAACGCTTAGATATTCCTCTGCGAAACAGCTACG 732
Db 2808 ATTTGATCTGAAAAAGCCTATGCGAAACGCTTAGACATTTGCTTCTGGAACCGCTGTGCG 2867
QY 733 CATTTGGGCGAGCAACACCGCAAGTGCAGTTGATTCCTCTTTGGTGGCGAGTAAAAAGT 792
Db 2868 CTTTGAACCGGTGAGAAAAAACCCTGGAAATTTGATTCAAATTTGGCGGTAAACCAACGCAT 2927
QY 793 GATTGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAAAGCCATATAAACATAAAGCGCT 852
Db 2928 TTACGGCTTTAACTCTCTTTGTGATCGCCAGCCGATACTGATGGCAAAAAAATCTGCTCT 2987
QY 853 TGACAAAGCGAAATCTCAAGGATTT-----ATCAAG 883
Db 2988 CAAACGCGCAAGAACAATGGCTTTGGTGTGTGAATTTGGGTTCGATATAAAAAATAAGG 3047
QY 884 TAAGGAGACTCCCATGAAA---ATGAAAAAACAAGATATGTAATACCTACGAGCCAC 940
Db 3048 AAAGGCAATTCGATGAAAAAATCTCTCGAAAAAGAAATATGTTTCTATGATGAGGCCAC 3107
QY 941 CAAAGCGCATAAAGTCGCTTAGAGATACCGATCTTTGGGCGAGAGTAGAACATGACTA 1000
Db 3108 TAGCGGCGAATAAGTAGATTTGGCGCATACCGACCTGATCTTAGAAGTCGAATGACTG 3167
QY 1001 TACCACCTATGGCGAAGAACTTAAATTTTGGCGGGGTAAACCTATCCGTGAGGGTATGGG 1060
Db 3168 CACCACCTATGGCGAAGAAATTAAGTTTGGTGGCGGTAAACCACTTCGCGATGGGATGGC 3227
QY 1061 TCAGAGCAATAGCCCTGATGA AAAACACCTTAGATTTTAGTCATCACTAAACGCGATGATAT 1120
Db 3228 ACAAAACAAGCCCGCAGCAGCCACCAACTCGATCTTTGTCTCACTAAACGCCCTGATCGT 3287
QY 1121 CGACTACACGGGATTTTACAAAGCCACATTTGGGATTA AAAACGCGAAATCCATGGCAT 1180
Db 3288 GGAATACACGGGATTTTAAAGCCGATATTTGGCATTTAA AAAATGGCAAAATCCATGGCAT 3347
QY 1181 TGGCAAGGCGAGAAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGGTCTGGGTGT 1240
Db 3348 TGGCAAGCAGGCNATAAGACATGCAAGATGGCGTTTGCAACANCTTTTGGTGGGCC 3407
QY 1241 GGGCACAAGACATAGCAGGGGAAGGTATGATTTATPACCGCTGGGGAAATCGATTCACA 1300
Db 3408 TGCTACTGAGGCTTTGGCCCTGAAGGGCTGATTTGTACAGCTGGTGGGATTGACACCCA 3467
QY 1301 CACCCACTTCTTTCTCCACAAACAATTCCTACCGCTCTAGCCAAATGGCGGTTCAACCAT 1360
Db 3468 CATCCACTTTATTTCTTCCCAACAAATCCCCACAGCATTTTGGCAGCGGGATCAACCAT 3527
QY 1361 GTTTGAGGCGGCACAGGTCCTGTAGATGGCAGCATGCGACTACTATCACTCCGGGCAA 1420
Db 3528 GATTGTGGGGNACAGGTTCCAGCTGATGGGACTTAACGCGACTACCATCACTCCGGGGCG 3587
QY 1421 ATGGAACCTTGCACCGCATGTTGGCGCAGCAGAGAGTATTTCTATGAATGTGGGCTTTTT 1480
Db 3588 CTGGAACTTAAACCACTGCTCGCTCTCTGGAAGATATGCCATGAACTTGGGCTATTT 3647
QY 1481 GGGCAAGGCANATAGCTCTAGCAAAAAAACAATTGTAGAAACAAGTAGAAGCGGCGCGAT 1540
Db 3648 GGGTAAAGGGAATGTGTCTTATGAACCCCTCCCTGGTCGATCAACTCGAAGCTGGAGCCAT 3707
QY 1541 TGGTTTTAAATTGCAATGAAGACTGGGGCAACAACCAAGTGGCGATCGATCACTGCTTGAG 1600

Db 3708 TGGCTTTAAATCCAGGAAGACTGGGGTAGCACCTTCGACGCCATCTACCAATTGCTTGA 3767
QY 1601 CGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATAGGCGAGG 1660
Db 3768 TGTGGCTGCAAAATACATGATGTGCAAGTGGCTATCCACACCGATACCTTGAATGAAGCGG 3827
QY 1661 TTATGTAGATGACACCCCTAAATGCAATGAAGCGGCGGCCATCCATGCTACCACTTGA 1720
Db 3828 CTGTGTGAAGACACTTTTGCAAGCCAATTGCTGGSCGCACTATCCACACTTTCCACACTGA 3887
QY 1721 GGAAGCGGTTGAAGGACACTCACTGATGTTATCACTATGGCAGGCGAGCTCAATATTTCT 1780
Db 3888 AGTGTCTGTGGGGGCACTGCTCGGATGTCAATTAAGATGTCTGGCGAAATTTAAACATCCT 3947
QY 1781 ACCTCTCTCCACACCCCACTATTCCTATACCATTAATACGTTGCGAGAACACTTGA 1840
Db 3948 CCAGCTTTCTACCAACCCCACTATTCCTTTACCGTGAATAACAAGCCGAACACATGGA 4007
QY 1841 CATGCTCATGACATGCCACCACTTAGACAAACGATCCGCGAGGATTTTACAATTTTCTCA 1900
Db 4008 CATGTTGATGTTGTCACCACTTGGATAAAAAATCAAGAAAGATGTCTCAGTTTGCTGA 4067
QY 1901 AAGCGGTATCGCCCGCGCTCTATCGCGCTGAAAGATGTCTCCATGATATGGGTGTGAT 1960
Db 4068 TTCTAGGATTTGGCCCGCAACCACTCGCGCTGAGGACAAACTCCACGATATGGGATTTT 4127
QY 1961 CGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCTGCGAGCGAAGTATTCCTCGAAC 2020
Db 4128 CTCTATCACCAGCTCTGACTCCCAAGCGATGGGCGGTGTAGGCGAGGTCATCACCCGAC 4187
QY 2021 TTGGCAGACTCGCGATGAAGATAAAAAAGAAATTTGGTAAAGCTTCTCTAAGATGGCAAGA 2080
Db 4188 TTGGCAAAACAGCGGCAAAAACAAAAGAAATTTGGTCTGCTGAGGAAAAGGCGA 4247
QY 2081 TAAAGATAATTTCCGATTAAGCGCTACATCTCCAAATACATATCAACCCCGCTTTGAC 2140
Db 4248 TAATGACAACTTCGCGCATCAAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGC 4307
QY 2141 CCAGCGGTGAGCGAGTATATCGCTCTGTGGAAGAGGGCAAGATCGCGAGCTTGGTGTG 2200
Db 4308 ACAGCGATTTCTGAAATATGTGCGCTCTGTAGAAAGTGGGCAAAATTCGCGGATTTGGTGTCT 4367
QY 2201 GTGGAATCTGCGCTTTTTCGCGTAAAAACCAAAATCGTGATCAAAAGCGGTATGGTGTG 2260
Db 4368 TTGAGTCTCGCTTTCTTTGGCATTTAAACCCACATGATCATAAAGCGGATTCATCGC 4427
QY 2261 CTTCTCTGAAATGGCGGATTTCTAAACGGCTCTGTGCCCACTCCCAACCGGTTTATTACCG 2320
Db 4428 ACTTTCTCAATGGCGCATGCGCATGCTCTATCCCACTCCCAACCGGTATTACCG 4487
QY 2321 CGAAATGTTTGGGATCACGGCAAGGCGAAATTTGACACCGATCACTTTTGTGTTCCAA 2380
Db 4488 CGAAATGTTTGGCACCACATGTTAAAGCCAAATTTGACACCAATATCACTTTGTATCCCA 4547
QY 2381 AGTCGCTATGAAATGGCGTGAAGAAAGCTCGGCTTAGAGCGCAAGTCTTACCGGT 2440
Db 4548 AGTGGCTTATGACAAACGGCATTTAAAGAAAGTGTGGGCTTGAAGAGTGGTTTTCCAGT 4607
QY 2441 CAAAACTGCGGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAAAACGGCAAAAT 2500
Db 4608 TAAAACTGCGCAACATCAACAAAAAGACTCAATTTCAACGATGTTACGCAACAT 4667
QY 2501 CACGTCGATCCGAAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTTAAACC 2560
Db 4668 CGAAGTCAATCTCTGAAACCTTACAAAGTTAAAGTGTGATGGCAAGAGGTTACTTCCAAAGC 4727
QY 2561 CACCTCCAGTGTCTTAGCCCGAGGCTACACTTTTCTTAGG 2604
Db 4728 AGCGGATAAAATAGCCTTAGCACAACTCTACAACCTGTGTTCTTAGG 4771

```

HECUREASE
LOCUS       2664 bp      DNA      linear      BCT 03-MAY-1995
DEFINITION Helicobacter heilmannii urease, complete cds's.
ACCESSION   L25079
VERSION     L25079.1  GI:793908
KEYWORDS    urease.
SOURCE      Candidatus Helicobacter heilmannii
            Candidatus Helicobacter heilmannii
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Helicobacteraceae; Helicobacter.
REFERENCE   1 (bases 1 to 2664)
AUTHORS     Solnick,J.V., O'Rourke,J., Lee,A. and Tompkins,L.S.
TITLE       Molecular analysis of urease genes from a newly identified
            uncultured species of Helicobacter
JOURNAL     Infect. Immun. 62 (5), 1631-1638 (1994)
PUBMED      8168924
COMMENT     On May 3, 1995 this sequence version replaced gi:529422.
            Original source text: Helicobacter heilmannii (individual_isolate
            2) DNA.

FEATURES             source
     source
1..2664
    /organism="Candidatus Helicobacter heilmannii"
    /mol_type="genomic DNA"
    /isolate="2"
    /db_xref="taxon:35817"
    /lab_host="BALB/c"
    200..204
    211..915
    /codon_start=1
    /transl_table=11
    /product="urease"
    /protein_id="AA065722.1"
    /db_xref="GI:529421"
    /translation="MKLTPKELDKMLHYAGELAKORKAGIKLNTVEAVALLSAHVH
    EARAGKQVADLMQGRLLKADDDVMPGVAMHIEVGI EAGPDGKLVTHIHPVEA
    GSDKLAGPVILKNEIDITLNAGHAVQLKYNKGRDPVQVGSHPF FEYNKLLDFPRE
    KAYKRLDISAGTVAFPEGEKTEVELIDIGGNKRIYGFNALVDROADHDKKLLAKR
    AKKHFGTINCQDNK"
    919..923
    930..2636
    /codon_start=1
    /transl_table=11
    /product="urease"
    /protein_id="AA065723.1"
    /db_xref="GI:529424"
    /translation="MKKISRKEYVSMYPTGDKVRLGDDTLLEVEHDCITYGEIHK
    FGKGTIRDMGNTSPSHSHELDVITNALIVDTGIYKADIGIKNGKIHGKAGNK
    DLQDGVNRLCVPAEALAAEGLIVTAGIDTHIFISPOQIPTAFASGITWIMGG
    TGPADGTNATTITPGRWNLKEMLRASEEYAMNIGYLGKGNVSEPALIDOLEAGATGF
    KIHEDWGSPSAINHALNTADKYDQVAIHTDTLNEAGCVEDTLEALAGRTIHTFHE
    GAGGHAPDVIKWAGFNILIPASTNETIPIPTKNTAEHMDMLCHLDLKNIKEDVFA
    DSRIRPOTIAEDKLDHMGIFISITSDSQAMRGVEVITRTWTQADKNKKKFGRLPEE
    KGDNDFRIKRYISKYITPAITHGIS EYVGSVEGVKYLADLWLSPAFFGIKENMIK
    GGFIALSOMGDANAS IPTQPVYRMPFHGHGKAKDTNITFVSQVAYENGIKHELGL
    QRVLFPVKCRNITKKDLKENDVTAHIEVNPETYKVKVDNGEVTSHAADKLSLAQLYN
    Lf"

RBS
CDS

RBS
CDS

Query Match      39.5%; Score 1138; DB 1; Length 2664;
Best Local Similarity 66.4%; Pred. No. 6.8e-217;
Matches 1709; Conservative 1; Mismatches 826; Indels 36; Gaps 4;

QY      63  TTGTTAATRCATATTATTTTAAATAATTACTTATATCATATATATATATATATTT 122
DB      70  TTGTGAATAAGATTAAACAAAGATAATAAT--ATTATCTTCGATAACAAATAATTAAC 127
QY      123  ACTTATATTAAGAGTTAATAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 182
DB      128  AACCATTAGCAACAAAGATACATAAATCTGCCTCGTTGATAGCTTGCCTATTTCAATA 187
QY      183  AAATTTAAACAGAGGAGTAAATAGTGAACCTACACCCAAAGAGCAAGAAAGTTCTTGT 242
DB      188  CCATTTTATACGAGGAGTTAGATGAACTGACACCTCAAGAGTTGGATAGTTGATGTC 247

```

```

243  TATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGAGCAGAGGCTTAAAGCTCAACCAAC 302
248  TCCATTATGCGGCGAAGTGGCTAGCAAAACAACGCAAGCAAAAGCAATTAAAGCTTAAATCA 307
303  CCGAAGCCATTGCTTACATTAGTCCCATATTATGACGAAGCGCGCGTGGCAAAAAA 362
308  CCGAAGCTGTAGCACTCATTTAGCGCCCATGTCATGGAAGAGCGCCGCGCGGTGAAAAA 367
363  CCGTTGCGCAGCTTATGGAAGAGTGCATGCACATTTTGAAGAAAAAGATGAAGTAATGCCCG 422
368  GCGTGGCGGATTGATGCAAGAGCGCAGCACTTACTCAAGCCGATGATGTGATGCTG 427
423  GGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAACCACTTTTCTCTGATGTGTAGCAAA 482
428  GCGTGGCGCATATGATCCACGAAGTGGGATTTGAAGCGGCTTTCCCGATGGGACAAAT 487
483  TTGTAATGTGAATTGGCCCATCGAACAGATGAGCACTTCAAGCGGGCGAGTGAAT 542
488  TAGTGACCATCCACACTCCCGTAGAAGCTGCGCAGCAAGCTTCTCCCGTGAAGTGA 547
543  TTGTTTCGATAAAGACATCGAGCTCAATGCAAGCAAGAAAGTAACCGAACTTTGAGTTTA 602
548  TCCTCAAAAAAGCAAGACATCACCTCAACGCCGCAAAACACGCCGTCCATTTAAAAGTCA 607
603  CTAATGAAGGCGCTAAATTCCTTGCATGTGGGTAGCCATTTCCACATTTCTTTGAAGCTAAC 662
608  AAAAAGGCGATCGCCCGTACAAGTGGGTTCACACTTCACATTTCTTTGAAGTGAATA 667
663  AGGCATTAATAATTCGATGTGAAAAAGCCTATGGAAGCGCTAGATATTCCTCTTGGCA 722
668  AGCTTTTAGACTTCGATCGCGCAAAAGCCTATGCAAAAGCGCTAGACATTTGCTTCTGGCA 727
723  ACAGCTACGCAATGGGGCAGGACAAACCCGCAAGTCAGTTGATTCCTCTTGGTGGCA 782
728  CCGCTGTGCGCTTTGAACCTGGGGAGAAAAACCGTGGAACTCATCGACATCGGTGGCA 787
783  GTAAAAAGTGATTGGCATGAACGGGCTTGTGAATAAATCAATCGCGATG----- 830
788  ATAAACGCAATTATGTTTCAACGCTCTAGTCGATCGCAAGCCGATCACGATGGCAAAA 847
831  -----AACGCCATAAACATAAAGCGCTTGACAAGCGGAAATCTCAACGATTTA 878
848  AACTCGCTTAAACGCGCTAAAGAAAAACACTTTGGCACTATCAACTGCGGTGCGACA 907
879  TCAAGTAAGGAGA-----CTCCATGAAATGAAAAACAAGAAATATGTAATACTTA 931
908  ACAAATAGGAAGAAAGAACCGGATGAAAAAATTTCTCGAAAAAGAAATATGTTTCTATGTA 967
932  CGGACCCCAAAAGCGCATAAAGTGGCTTAGGAGATACCGATCTTTGGCGAGAGTAGA 991
968  TGGACCCCACTACGGGCGATAAAGTGGCTTGGCGAGACTGACCTGATCTCTAGAGTCA 1027
992  ACATGACTATACCACTATGCGGAAGAACTTAAATTTGCGCGGGTAAAAAATATCCGTGA 1051
1028  ACATGACTGACCACTTATGGCGAAGAAATCAATTCGGGGCGGAAAAACCATCCGCGA 1087
1052  GGGTATGGGTACAGCAATAGCCCTGATGAAAAACACCTAGATTTAGTCACTCACTAACGC 1111
1088  TGGGATGGGACAAACCAACAGCCCCAGCAGCCATGAGCTCGATCTTGTCAATCACCACGC 1147
1112  GATGATTTATCGACTACACCGGATTTACAAAGCCGACATTCGGATTTAAAAACGGCAAAAT 1171
1148  GCTGATTTGTGATTTACACCGGCATTTACAAAGCCGACATTCGGCATTTAAAAACGGCAAAAT 1207
1172  CCATGGCATTTGGCAAGCAGGAAAAAAGGACATGCAAGATGCGGTAAAGCCCTCATATGCT 1231
1208  CCACGGGATTTGGCAAGCCGCAACAAAGACTTCAAGATGGCGTTTTCGCAACAGGCTCTG 1267
1232  CGTGGGTGTGGGCAAGAGCACTAGCAGGGGAAGGTATGATTTATTAACGCTGGGGGAAT 1291
1268  CGTAGGACCTGCTACAGAGCTCTCGCTGCTGGAAGGCTTGATTTGTTACCGCGGTGGGAT 1327

```

QY 1292 CGATTACACACCCACTTCTCTTCTCCACAAACAATTCCTACCGCTCTAGCCAAATGGCGT 1351
Db 1328 CGACACCCACATCCACTTCTATTTCTCTCAACAAATCCCTACTGCTCTTGTAGCGGAAT 1387
QY 1352 TACACACCATGTTTGGAGCGGCACAGGTCTCTAGATGCGACGAATGCGACTACTATCAC 1411
Db 1388 CACCAACCATGATCGGGGGCGGACTGGACCTGAGATGCGACCAACGCCACCAACCATCAC 1447
QY 1412 TCCGGGCAAAATGGAACTTCACCGCATGTTGGCGCGCAGCAGAGATATTTCTATGAATGT 1471
Db 1448 TCCCGCGCGCTGGAATCTTAAAGAAATGCTCGCGCTTCTGAAGATACGCCATGAACCT 1507
QY 1472 GGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTTGTAGAACAAAGTAGAAGC 1531
Db 1508 TGGCTACCTTGGTAAGGGGAATGTGCTTTTGAACCTGCTCTCATTTGACCGAGCTAGAAGC 1567
QY 1532 GGGCGCGATGCTTTTAAATTCGATCAAGACTGGGGCACAACACCAAGTGGCGATCATCA 1591
Db 1568 CGGCGCGATTTGGCTTTTAAATCCACGAAGACTGGGGCAGCACACCTTCAGCCATCAACCA 1627
QY 1592 CTGCTTGAGCGTGGCAGATGAATACCATGTGCAAGTTTGTATCCACACGATACAGTCAA 1651
Db 1628 CGCTCTAAACATCGCTGAACAATACGATGTGCAAGTGGCGATCCACACGACACCTTGA 1687
QY 1652 TGAGGCAGGTTATGTAGATGACACCTTAATGAATGAACGGGCGGCCATCCATCCCTA 1711
Db 1688 TGAAGCGGCTGTGTGAAGACACCTTAGAAGCATCGCTGGAGCGACCATCCACACCTT 1747
QY 1712 CCACATTTAGGAGCGGGTGGAGGACACTCACTGATGTTATCACCATGCGAGCGAGCT 1771
Db 1748 CCACACCGAAGTGTCTGGGCGGACACGCTCTGACGATCAAAATGGCGGCGAAAT 1807
QY 1772 CAATATTTACCTCTCCACACCCCTACTATTTCCCTATACCATTAATACGGTTCGAGA 1831
Db 1808 TAACTCTCTCTGCTTTACCAACCCACCATCCCTTTCCACAAAAACACAGAAGCCGA 1867
QY 1832 ACACCTTAGACATGCTCATGACATGCCACACCTTAGACAAACGATCCGGAGGATTTACA 1891
Db 1868 ACACATGGACATGCTTATG--TGCCACCATTTGGATAAAAAACATCAAGAAGATGGGA 1924
QY 1892 ATTTCTCAAAGCGTATCGGCGCGCTCTATCGCGCTGAAGATGTGCTCCATGATAT 1951
Db 1925 ATTTGCGGACTCACGATATCGGCGCGCAACCAATTTGGCGGGAAGATAAATCCACGACAT 1984
QY 1952 GGGTGTATCGGATCAGACATCGATTCGCAAGCAATGGGCGGTCGAGGCGCAAGTAT 2011
Db 1985 GGGGATTTTCTCATCACTAGCTCCGACTCTCAAGCCATGGGTGCGGTGGGCGAAGTAT 2044
QY 2012 TCTCGAACTTGGCAGACTGGGATGAAGATAAAAAAGAAATTTGGTAAGCTTCTCTGAAGA 2071
Db 2045 CACCGCGACTTGGCAACAGCGGACAAAAAACAANAAGAAATTTGGCGCTTGTCTGAGA 2104
QY 2072 TGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCAAAATACATATCAACCC 2131
Db 2105 AAAAGCGGACACGACAACTTCGCGCATCAACGCTATTTTCCAAATACACCATCAACCC 2164
QY 2132 CGCTTTGACCCAGCGGTGAGCAGTATATCGCTCTGTGGAGAGGGGCAAGATCGCGGA 2191
Db 2165 TGGCATCACACAGCGGATTTCTGAATATGTGCGCTCTGTGAAGATGGGTAATACGCTGA 2224
QY 2192 CTGTGGTGTGTGAATCTCGCTTTTGGCGTAAAAACCCAAATCGTGTATCAAAAGCGG 2251
Db 2225 CTGTGGTGTGTGGAGCGCTCGCTTTCTTGGCATTTAAACCCCAACATGATCATCAAGCGG 2284
QY 2252 TATGTGTGCTTCTCTGAAATGGGCGGATTTAAACGGGTCTGTGCCCATCTCCCAACCGGT 2311
Db 2285 TTTTCATTGCGCTTTCTCAAAATGGCGGATGCAACGCTTCTATCCCACTCCCAACCGGT 2344
QY 2312 TTATTAACCGGAAATGTTTGGGATACAGGCAAGGCGAAATTTGACACAGCATCACTTT 2371
Db 2345 GTACTACCGCGAAATGTTTGGGCGACCAACGCGCAAGGCGAAATTTTGAACCAACATCACTTT 2404
QY 2372 TGTTTCCAAAGTCGCTATGAAATGCGGTGAAGAAAAAGCTGGGCTTAGAGCGCAAGT 2431

Db 2405 CGTGTCTCAAGTGGCTTATGAAAAACGGCATTAACACACGAGTTGGCTTGCAAAAGATTGT 2464
QY 2432 TCTACCGGTCAAAAACATGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGACAAAAC 2491
Db 2465 GTTGTCTGTGAAAAACTGCGCGCAACATCACAAAAAAGACCTTAAGTTCAACGATGTAC 2524
QY 2492 GGCAAAAATCACCGTTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGCAAACTCTGCAC 2551
Db 2525 CGCACACATCGAAGTCAACCTGAAACCTACAAAGTGAAGTGGATGGACGAGATTAC 2584
QY 2552 CTCTAAACCCACCTGCGAAGTCCCTAGCCAGCGGTACACATTTCTTCTTAG 2603
Db 2585 CTCCCATGCGGCTGACAAATTTGACCTAGCACACTCTATAACCTATTCTTAG 2636

RESULT 13
AE001446/c
LOCUS
DEFINITION
Helicobacter pylori, strain J99 section 7 of 132 of the complete genome.
ACCESSION
AE001446 AE001439
VERSION
AE001446.1 GI:4154573
KEYWORDS
SOURCE
ORGANISM
Helicobacter pylori J99
Helicobacter pylori J99
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.
REFERENCE
1 (bases 1 to 12037)
Alm, R.A., Ling, L.S., Moir, D.T., King, B.L., Brown, E.D., Doig, P.C., Smith, D.R., Noonan, B., Guild, B.C., deJonge, B.L., Carmel, G., Tummino, P.J., Caruso, A., Uria-Nickelsen, M., Mills, D.M., Ives, C., Gibson, R., Merberg, D., Mills, S.D., Jiang, Q., Taylor, D.E., Vovis, G.P. and Trust, T.J.
Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
Nature 397 (6715), 176-180 (1999)
9923682
2 (bases 1 to 12037)
King, B.L., Alm, R.A. and Trust, T.J.
Direct Submission
Submitted (12-JAN-1999) Astra Research Center Boston, 128 Sidney Street, Cambridge, MA 02139, USA
Address all correspondence to: hp@arcb.us.astro.com or Richard A. Alm, Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Lo-See L. Ling, Donald T. Moir, Douglas R. Smith, Rene Gibson, and Gerald F. Vovis are with Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA, 02453. Qin Jiang and Diane E. Taylor are with the University of Alberta Department of Medical Microbiology and Immunology, Edmonton, Alberta, Canada, T6G 2H7 and the Canadian Bacterial Diseases Network. All other authors are with Astra Research Center Boston, 128 Sidney Street, Cambridge, MA, 02139. Putative identifications, sequence alignments, and name and sequence search capability are available at ARCB's World Wide Web site. (URL: http://www.astro-boston.com/hpylori).
FEATURES
source
1. 12037
/organism="Helicobacter pylori J99"
/mol_type="genomic DNA"
/strain="J99"
/db_xref="taxon:85963"
complement(143. 1852)
/gene="ureB"
/note="synonym: jhp0067"
complement(143. 1852)
/gene="ureB"
/note="similar to H. pylori 26695 gene HP0072"
/codon_start=1
/transl_table=11
/product="UREASE BETA SUBUNIT"
/protein_id="AAD05651.1"

```
/db_xref="GI:4154576"
/translation="MKKISRKEYVSMYGPPTTGDVKRLGDDTLIAEVHEHYTIYGEELK
FGGKTLREGMSNNPSKEELDIIITNALIVDTYGIYKADIGIKDGGKIAGIGGKNG
TMDQGVKNNTISVPATLALAGEGLIVTAGGIDTHIFISPOQIIPAFASGVTTMIGGG
TGPDAGTNATTIPGRNLKWLRAAEYSNNLGLAKGNASNDASLADQIEAGAIGF
KIHEDMGTSPSANHALDVADKYDVOVAIHDTDTLNEAGCCVEDTAAIAGTMTHTFHE
GAGGCHAPDIIKVAGENHILPASTNPPIPTVWTEAEHMDMLVCHHLDKSIEKDVQF
ADSRIPQTIAAEDTLHDMGIFISITSSDQAMGRVGEVITRTWQTDARKKKEFGRLIKE
EGKDNDFRIKRYLSKTYTINPAIHGISEYVSGVKGADVLWSFAFGVKFNMI
KGQFIALSOMGDANASIPTPQPVYVREMFHAHGKAKYDANITFVSQAAYDKGIEBELG
LERQVLPVKNCNRTTKDMQFNDTTAHEIVNPETYHYFVVDGKEVTSKPANKVSLAQLF
SIP"
complement(1856. .2572)
/gene="ureA"
/notes="synonym: jhp0068"
CDS
complement(1856. .2572)
/gene="ureA"
/notes="similar to H. pylori 26695 gene HP0073"
/codon_start=1
/transl_table=11
/product="URBASE ALPHA SUBUNIT"
/protein_id="AAD05652.1"
/db_xref="GI:4154577"
/translation="MKUTPKELDKMLHYAGELARKKKEKGIKLNYYEAVALISAHIM
EEARAGKTTAAELMQEGRTLLKPDNDVMDGVASMTHEVGIEAMFDGTLKLVHTPIEA
NGKLVGELEFLKNEDITINEGKAVSVKXNVGDRPVQIGSHFHF FEVNRCLDPDREK
TFGKRLDIAAGTAVPEPGEKSVELIDIGNRRIFGFNALVDQRADNESKIALHRA
KERGFHAKSDNNVKTIKE"
complement(2926. .3417)
/gene="lspA"
/notes="synonym: jhp0069"
CDS
complement(2926. .3417)
/gene="lspA"
/notes="similar to H. pylori 26695 gene HP0074"
/codon_start=1
/transl_table=11
/product="LIPOPROTEIN SIGNAL PEPTIDASE"
/protein_id="AAD05653.1"
/db_xref="GI:4154578"
/translation="MLNTTQKSLLVFMGVFFLIPGVDOAKHAILEGFHYESLVIDIV
LVNRQVAFSLFEGGLYQIILLIGLFIPLMRQKELFKSHAIFFGMVFGAGVSN
VLDRFVHGVDVYVYHYGDFDAIFNPAFVMDVGVGVLRLRQFFKQKQKIKGIIT
LRF"
complement(3411. .4748)
/gene="glmM"
/notes="synonym: jhp0070"
CDS
complement(3411. .4748)
/gene="glmM"
/notes="similar to H. pylori 26695 gene HP0075"
/codon_start=1
/transl_table=11
/product="phosphoglucosamine mutase"
/protein_id="AAD05654.1"
/db_xref="GI:4154579"
/translation="MKIFGTDGVRGKAGVKLTPMFVMRLGIAAGLYFKGHSOTNKILLI
GKDRKSGVMVNALYSALTSIGVNIQIGMPPTPAIAFLTEDMRCDAGIMISASHNP
FEDNGKFNFSYGYKLUKEBERAIEEIHFEELLHSSYKYGESVGSAKRIDDDVIGRYI
AHLKHSFPKHLNLOKURIVLDTANGAAKYAPVVVSELGADVLVINDEPNGCNINEQC
GALHPNQLSDEVYKLRADLGFADFDADRLVVDNLGNIHVGDKLGLVGLYQKSKNA
LSSQAVVATSMNLAKLEYLKSQDLELHCAIGDKFVSECWRLNKNANFGGQSGHIIF
SDYAKTGIDGVALOVSALVLESKLVSIIALNPPELYPSQILVNLNIOKKPPELSLKG
SALLKELDQLEIRHLIRYSTENKLRILLKAEKLEKSKQELKEFFEQQLC"
4838. .5107
/gene="rpsT"
/notes="synonym: jhp0071"
CDS
4838. .5107
/gene="rpsT"
/notes="similar to H. pylori 26695 gene HP0076"
/codon_start=1
/transl_table=11
/product="30S RIBOSOMAL PROTEIN S20"
/protein_id="AAD05655.1"
/db_xref="GI:4154580
```

```
gene
/translation="MANHKSAEKRIQTIKTERNRFYKTKVKNIIKAVREAVAINDV
AKQRRLKIANKELHKFVSKGLKKNNTASRKSRLNASVKKIALA"
5226. .6284
/gene="prfA"
/notes="synonym: jhp0072"
CDS
5226. .6284
/gene="prfA"
/notes="similar to H. pylori 26695 gene HP0077"
/codon_start=1
/transl_table=11
/product="PEPTIDE CHAIN RELEASE FACTOR 1"
/protein_id="AAD05656.1"
/db_xref="GI:4154581"
/translation="MSIAEKLSLILKRYDELTAALLSSVEVVDIKKLTLSKQSESI
BEISVASKEYLSVLGEEKENKELSELKELKLEIKSELETAIKQLLIPK
DPNDKNIYLELRAGTGDEAGIFVGDLPFKAYCRYADLKKWKVEIVSSSENSVGYKE
IIVLIKGVYSRLKFEAGTHRVQRPETESQGRITSAITVAIMEVDVDEVSINPS
DLKIEVFRAGGHGGCCVNTTDSAVRITHLPNIVSVMODEKSHQKNDKALKILKARL
YEKIEEQOLANAKRKQGVSGDRSERIRTYNPQNRLSEHRINLTLSLEEMLSG
NLDEVINPLIAHAQSQFE"
6698. .7465
/gene="jhp0073"
6698. .7465
/gene="jhp0073"
/notes="similar to H. pylori 26695 gene HP0078 and HP0079"
/codon_start=1
/transl_table=11
/product="putative Outer membrane protein"
/protein_id="AAD05657.1"
/db_xref="GI:4154582"
/translation="MKKIFLGMALFVSMAEKSGAFLLGGGFQYSNLENQNTTRTPSA
NNNTPINTSMFGNNQAPQAPETPSVINNTNYQMGYGDAMAGYKNFPGTKRFGFRY
GYSYNHAANLSFVSGKLGIMDGASOVNNTYGVGFDALYNFYESKEGYNTAGLFGVGF
LGDSFIYQGESYLSKQICNNTAGCSASMTSYFQMPVEFGFRSNFSKSHSIEGVF
KLPLFTNQFYKERGVDSVDVFKRNFISIYFNYMINL"
7716. .9488
/gene="jhp0074"
7716. .9488
/gene="jhp0074"
/notes="similar to H. pylori 26695 gene HP0080"
/codon_start=1
/transl_table=11
/product="putative"
/protein_id="AAD05649.1"
/db_xref="GI:4154574"
/translation="MKAIKILLIWLISNAISVNEALPDLKDSQLKGELTPKIVDFGG
YKSNTEWGATALNTYINANGDAKFSALVERKRFNSGILGNFRAHRAHRLQALQKN
LYKCLIKIIRDSFYSYRTGIYPLGILSKDQKTAQLADLSVVGAYLKKQKQENKAQ
SPYRSNNYNSYTSPPYGYMGYMGYMGYMGYDFYDFYDGYGYPNNFPMQ
VODYLMLNMYALDQEBILDHASINOLDPTDDDDKDDKSSQAPANLMSFYRDPK
PSKDIQTNRLSALVNLNSHMLKDNSLPHTKAMPTKSDAITSOAKELNHLVGOIKE
MKQGASPNKIDSVVNKAMEVRDKLDNLNLQDLNDLKOQGLSSEQQAQVDKALDSVQ
QLSHSSDVVNGYLDGSLKIDGDRDDLNDAINNPMQQAQPTPINMNDHANDSKDQ
GGNALINPNATNDHDDHMDTNTDTSNANDPTDDKASGNTGDMNNTDTGTND
TGNTDTGTNDMMNNGNDDTGTNTDDMGNSNDMGDDMNNANDMNDMDGNSDDMDG
MGDMDMGDMGDMGDMGDMGN"
complement(9912. .11933)
/gene="jhp0075"
complement(9912. .11933)
/gene="jhp0075"
/notes="similar to H. pylori 26695 gene HP0082"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein (MCP)"
/protein_id="AAD05650.1"
/db_xref="GI:4154575"
/translation="MKSTRIGSKIYMWYCAVVIVISAVMGVIISYKVESVLOSQATEL
LQKAQLVSYFKIQIKMGIEMGANTLERLSDENGAINDTLKRRLMSEFLLANPHVLL
YSAITNNNERMITAMNMDSKIAYENTALNENWTQIHSLSKITSRSDPYKYEVNGDKI
YGMDDITLPLMGKNQNVIGALNFFLNIDAFYTDVVGKKSNTLPMGDKRILLNPNREI
QDKILSAINPDKRAKAYEYVYQNEAGTSLSGNTETFLAIQDFDFEKGNGN
HWRWAIGKYVNSLVFKEATNTKFIITTLILGLVLVLAFLVFLVFIISNLITKRISRVN
```

[illegible]

Db	615	CAGCGATCGCTCATGAGGATTAGCGAGTATATAGAGTTCTGTAGAAAGTGGGCAAAAGTGCGTG	556	FEATURES	Location/Qualifiers
QY	2191	ACTTGGTGGTGTGGAATCCTGCCTTTTTCGGCGTAAACCCAAAATCGTGATCAAAAGCG	2250	source	1. .10903
Db	555	ACTTGGTATTTGGAGTCGAGCATCTTTGGCGTGAACCCCAACATCATCATCAAGGCG	496		/organism="Helicobacter pylori 26695"
QY	2251	GTATGGTGGTCTTCTCTGAATAGGCGGATTTAAACGGCTCTGTGCCCACTCCCCAACCGG	2310	gene	/mol_type="genomic DNA"
Db	495	GATTCAATGGCTTAAGCCAAATGGCGGATGCGAACGCTTCTATCCCTACCCCAACACCGG	436	CDS	/strain="26695"
QY	2311	TTTATTACCCGGAATGTTTGGGCATCAGCGACGGGAATTTGACACGAGTCATCTT	2370		/db_xref="taxon:85962"
Db	435	TTTATTACAGAGAAATGTTCCGCTCACCATGGTAAAGCTAATACGATGCAAAACATCACTT	376		/complement (127. .1836)
QY	2371	TTGTTTCCAAAGTCGCTATGAATAAGCGGTGAAGAAAGCTGGCTGTAGAGCGCCAAAG	2430		/complement (127. .1836)
Db	375	TTGTGTCCTAAGCGGCTTATGACAAAGGCATTAAGAAGAAATTAGGCGTTGAAGACAAG	316		/gene="HP0072"
QY	2431	TTCTACCGGTCAAAAACCTGCCGTAAACATCAACGAAGAAGCTTCAAGTTCAACGACAAA	2490		/note="similar to GB:M60398 SP:P14917 PID:149011
Db	315	TGTTGCCGTTAAAAATTGCAGAAACATCATCTAAAAAAGACATCAATTCATGACACTA	256		GB:AE000511 percent identity: 100.00; identified by sequence similarity; putative"
QY	2491	CGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCCTTTGTAGATGGCAAACTCTGCA	2550		/codon_start=1
Db	255	CCGTCACATTTGAAGTCAATCTCTGAAATTTACCATGTGTTCGTGGATGGCAAGAGTCA	196		/transl_table=11
QY	2551	CCTCTAAACCACTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTCTTAGG	2604		/product="urease beta subunit (urea amidohydrolase)
Db	195	CTTCTAAACGAGCTAATAAGTAGCTGGCTCAACTCTTTAGCATTTTCTTAGG	142	gene	(ureB)"
RESULT 14				CDS	/protein_id="AAD07143.1"
AE000529/c					/db_xref="GI:2313153"
LOCUS	AE000529	10903 bp	DNA	linear	BCT 06-APR-1999
DEFINITION	Helicobacter pylori 26695 section 7 of 134				of the complete genome.
ACCESSION	AE000529	AE000511			
VERSION	AE000529.1	GI:2313152			
KEYWORDS					
SOURCE	Helicobacter pylori 26695				
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.				
REFERENCE	1 (bases 1 to 10903)				
AUTHORS	Tomb,J.F., White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzgerald,L.M., Lee,N., Adams,M.D., Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R., Peterson,J.D., Kelley,J.M., Cotton,M.D., Weidman,J.M., Fujii,C., Bowman,C., Watthey,L., Wallin,E., Hayes,W.S., Borodovsky,M., Karp,P.D., Smith,H.O., Fraser,C.M. and Venter,J.C.				
TITLE	The complete genome sequence of the gastric pathogen Helicobacter pylori				
JOURNAL	Nature 388 (6642), 539-547 (1997)				
PUBMED	9252185				
REFERENCE	2 (bases 1 to 10903)				
AUTHORS	Tomb,J.F., White,O., Kerlavage,A.R., Clayton,R.A., Sutton,G.G., Fleischmann,R.D., Ketchum,K.A., Klenk,H.P., Gill,S., Dougherty,B.A., Nelson,K., Quackenbush,J., Zhou,L., Kirkness,E.F., Peterson,S., Loftus,B., Richardson,D., Dodson,R., Khalak,H.G., Glodek,A., McKenney,K., Fitzgerald,L.M., Lee,N., Adams,M.D., Hickey,E.K., Berg,D.E., Gocayne,J.D., Utterback,T.R., Peterson,J.D., Kelley,J.M., Cotton,M.D., Weidman,J.M., Fujii,C., Bowman,C., Watthey,L., Wallin,E., Hayes,W.S., Borodovsky,M., Karp,P.D., Smith,H.O., Fraser,C.M. and Venter,J.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	3 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	4 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	5 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	6 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	7 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	8 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	9 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	10 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	11 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	12 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	13 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	14 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	15 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	16 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	17 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	18 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	19 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	20 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	21 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	22 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	23 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	24 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	25 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	26 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	27 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	28 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	29 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	30 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	31 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	32 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	33 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	34 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	35 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	36 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	37 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	38 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	39 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	40 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	41 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	42 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	43 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	44 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	45 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				
REFERENCE	46 (bases 1 to 10903)				
AUTHORS	White,O.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
PUBMED	9252185				

```
/note="similar to GB:M60398 SP:P25177 PID:149008
GB:AB000511 PID:2313156 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="urease protein (ureC)"
/protein_id="AAD07146.1"
/db_xref="GI:2313156"
/translation="MKIFDTGVRGKAGVGLTPMFVRLGIAAGLYFKKHSOTNKILI
GDKTSGVMVNALVSALTSGYVNIQIGPMPPTAIAFLTEDMRCADGIMISASHNP
FDNGIKFNSGYKIKEREKAIEIFHDELLHSSYKVGSGSAKRIIDDVIGYI
AHLKSPFKHLNLSRLRVLDTANGAAYKAVPVFSELGADVLIINDEPGNCINQOC
GALHPNQLSQEVKKRADIGFAPDGDRLVVDNIGNIVHDKLIGLVIGVYOKSKNA
LSSQAVATMNLAKLEYLKSQDLKRCALKEFVSECMQLNKFNGEBSGHIIF
SDYAKTGDGLCALQVSLVLESQVSVLNPFEYPOSVLNINQVKPPLESLSKY
SALLKELDLERHLIRYSGTENKLRILLEANDEKLESQMQLKEFFEGHLC"
4796..5065
/genes="HP0076"
4796..5065
/genes="HP0076"
/note="similar to SP:P48948 PID:755068 percent identity:
41.38; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S20 (rps20)"
/protein_id="AAD07147.1"
/db_xref="GI:2313157"
/translation="MANHSAEKRIQTIKRTNRNFYTKIKNIKAVREAVAVNDV
AKAQRLEKIANKELHKFVSGILKNTASRKVSRNLASVKKIALA"
5189..6247
/genes="HP0077"
5189..6247
/genes="HP0077"
/note="similar to GB:M11519 SP:P07011 PID:147567
PID:968930 GB:U00096 percent identity: 52.87; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptide chain release factor RF-1 (prfa)"
/protein_id="AAD07148.1"
/db_xref="GI:2313158"
/translation="MSILAELKSSILKRYDELTAALSSAEVSVSDIKKLTSLKEQSSI
EPISTASKEYLVLENKENKELLEDELSELAKEIKLEIOKSLDETAIKOLLIPK
PDNDKNIYLELRAGTGGDEAGIFVGDLPKAYCRYADLKKWKEIIVSSSENSVGGYKE
ITALIKGVYSRLKEAGTHRVORVPETESQRIHTSAITVAIMEPVDVDSINPS
DLKIEVFRAGGGQCWNTDSAVRITHLPNTISVSMQDEKSHKDKALKILKARL
YKQIEEQULANAKRQVSGSDRSRIRITNYPQNRUSEHRINLTLSLEBIMLSG
NLDEVINPLIAHAQSQFE"
6658..6915
/genes="HP0078"
6658..6915
/genes="HP0078"
/note="hypothetical protein; identified by GeneMark;
putative"
/codon_start=1
/transl_table=11
/product="H. pylori predicted coding region HP0078"
/protein_id="AAD07149.1"
/db_xref="GI:2313159"
/translation="MKKVFLGMALAFSVSMAEKSAGFLGGFGVSNLENQNTTRTPGA
NNPTPIDTSMFGSNKTPAQETQSAKPDTKVNPSPASWMKK"
6926..8713
/genes="HP0079"
6926..8713
/genes="HP0079"
/note="similar to GB:AB000511 PID:2313160 percent
identity: 100.00; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein (omp3)"
/protein_id="AAD07150.1"
/db_xref="GI:2313160"
/translation="MKKSPKLGKLVLSAAGSVLLGSNNATDLETTAALQKSHVFCNY
AEKDQSKLTSPTQDDQKVAQNTASNDSEATTLENTASTNTTATTDETYTKST
DTTVAGAQQVETDNTAVQSAEQTKTDVAQVADAKDFEDTTFQADQAEQTAEK
ALQOAESKLTNDQOTLNTADQTKPTPTPTKBPKHTASSPSPAPKAD
ETSGTSPASGSSVASOLTKDITWNNLKSVSAMNTTLCGVETMSQOTATIGNLLNS
STDLSSVFNQAQGLNSAFSTLESQAQNTLKGYNSSSATIGQLTNGSNVAVGALDXAIN
QVDMALADLSAADTQKTQVTLATASPSITTDIAINFLNALKNSLMAQKDAFLNVHK
NIQTVAQAQETTYTPTVNTNNTQYGVDMAGYKWFPGTKRFPGRSYGYSYNHA
NLSPVQSOLGIMEGASQVNTYGVDFVLNIFYESKEGYNATGLFLGFLGGDSFIV
QGSYLKSOHMI CNNTAGCSASMTNTSYFQMPVERFGRSNFNSKHSIEVGFKLPLFTNQ
FYKERVGDGSDVDFYKFNFSIYFNYMINF"
8959..10740
/genes="HP0080"
8959..10740
Query Match 37.5%; Score 1081.6; DB 1; Length 10903;
Best Local Similarity 66.2%; Pred. No. 1e-205;
Matches 1626; Conservative 0; Mismatches 794; Indels 38; Gaps 3;
Qy 182 AAAATTTAACACAGAGTAAATAGTGAAACTACACCCCAAAGAGCAAGAAAAGTTCTTTG 241
Db 2580 ACACCTTTAAGATAGGAGAATAAGATGAACTCACCCCAAAGAGTTAGACAAGTTGATG 2521
Qy 242 TTATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAA 301
Db 2520 CTCCACTATGCTGAGAATTTGGCTTAAAAAAGCAAGAAAAGGCAATTAAGCTTAACAT 2461
Qy 302 CCGAAGCCATTGCTTACATTTAGTGCCCATATTATGACGAAGCGCGCTGGGAAAAAAA 361
Db 2460 GTAGAAGCGGTAGCTTTGATTGATGCCCCATATTATGAAGAAGCGGAGAGCTGGTAAAA 2401
Qy 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGTAAAAAAGATGAAGTAATGCC 421
Db 2400 ACTCGCGCTGATTGATGCAAGAGGGCGCACTCTTTTAAACCAGATGATGTGATGAT 2341
Qy 422 GGGGTGGGTAAATATGGTTCCGATCTAGGTGTAGAGCCACTTTTCTTCGTATGGTACGAAA 481
Db 2340 GCGGTGCGAAGCATGATCCATGAAGTGGGTATTGAAGCGATGTTTCTCGATGGGCAAAA 2281
Qy 482 CTTGTAAGTGTGANTTGGCCCATCGAACCATGAGCACTTCAAAGCGGCGAAGTGAA 541
Db 2280 CTGTAACCGGTGCATACCCCTATTGAGGCCAATGGTAAATAGTTCTCGGTGAGTTG--- 2224
Qy 542 TTTGGTTGCGATAAAGACATCGAGCTCAATCGACGCAAGAAAGTAACCGAACTTGAGGTT 601
Db 2223 TTCTTAAAAATGAAGACATCACTATCAAAGAAAGCAAAAAGCCGTAGCGTGAAGTT 2164
Qy 602 ACTAATGAAGGGCCCTAAATCCTTCATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAC 661
Db 2163 AAAAAATGTTGGCGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 2104
Qy 662 AAGGCATTAATAATTCGATCGTGAAGAAAGCCCTATGGCAAGCGCTAGATATTCCTCTGCG 721
Db 2103 AGATGCTTAGACTTTGACAGAGAAAAAACTTTTCGTTAAACGCTTAGACATTTGCGAGCGG 2044
Qy 722 AACACGCTAGCATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTGTTGTCG 781
Db 2043 ACACGGTAAAGTTTGAGCCCTGGCGAAGAAAAATCCGTAGATTGATTGACATTTGCGGT 1984
Qy 782 AGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACCGCATAAA 841
Db 1983 AACAGAGAATCTTTGGATTAAACGCGTTGTTGATAGCAAGCAGACACCAAGAACAAA 1924
Qy 842 CATAAAGCGCTTGACAAGGGCGAAATCTCACGGATTT----- 877
Db 1923 AAAATTGCTTTACACAGAGCTTAAAGAGCGTGGTGTTCATGGCGCTTAAAGCGATGACAAC 1864
Qy 878 -----ATCAAGTAAGGAGACTCCCATGAAA---ATGAAAAACAGAAATATGTAAAT 926
Db 1863 TATGTAAAAACAAATTAAGGAGTAAGAAATGAAAAAGATAGCAGAAAAAGAAATATGTTCT 1804
Qy 927 ACCTACGAGCCCAACCAAGGGCGATAAAGTGGCGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
```



```
Db      1803 ATGTATGGCCCTACTACAGCGGATAAAGTGAGATTGGCGATACAGACTTGATCGCTGAA 1744
Qy      987 GTAGAACATGACTATACCACCTATGGCGAAGAACTTAAATTTGGCGGGTAAACCTATC 1046
Db      1743 GTAGAACATGACTATACCACCTTATTTGGCGAAGAGCTTAAATTTGGTGGCGTAAACCCGTG 1684
Qy      1047 CGTAGGGTATGGGTAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCACTACT 1106
Db      1683 AGAGAAGGCATGAGCCCAATCCAAACACCTAGCAAGAAGAATTTGGATCTAATCATCACT 1624
Qy      1107 AACCGATGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTTAAACAGGC 1166
Db      1623 AACGCTTTAATCGTGGATATACCCGGTATTTATAAGCGGATATTTGGTATTTAAAGATGGC 1564
Qy      1167 AAAATCCATGGCATTTGCCAAGGAGGAAACAAGACATGCAAGATGGCGTAAAGCCCTCAT 1226
Db      1563 AAAATCGCTGGCATTTGTAAGGGGTAAACAAAGCATGCAAGATGGCGTTAAACACAT 1504
Qy      1227 ATGGTCGTGGGTGGGCACAGAACACTACTAGCGGGAAGGTATGATTTACCGCTGGG 1286
Db      1503 CTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCCGGTGAAGGTTTGATCGTAACTGCTGGT 1444
Qy      1287 GGAATCGATTACACACCCACTTCTCTCCACAAATTCCTACCGCTCTAGCCAT 1346
Db      1443 GGTATTGACACACATCTCACTTCATTTCAACCCCAACAAATCCCTACAGCTTTTGCAGC 1384
Qy      1347 GCGTTTACAACTATTTGGAGGGCGCACAGGTCCTGTAGATGCGACGAATGCCACTACT 1406
Db      1383 GGTGTAAACCAATGATTTGGTGGCGAACTGGTCTCTGTATGCACTAATTCGCACTACT 1324
Qy      1407 ATCACTCGGGGCAAAATGGAACTTTGCACCGCATGTTGGCGCAGCAGAGAGTATTTCTATG 1466
Db      1323 ATCACTCAGCGAAGAAATTTAAATTTGAATGCTCTAGAGCGGCTGAAGAATATTTCTATG 1264
Qy      1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAACAACTTTGTAGAACAGTA 1526
Db      1263 AACTTAGGTTTCTTTGGCTAAAGGTAAACGCTTTTAAACGACGCGAGCTTAGCGCATCAAAAT 1204
Qy      1527 GAAGCGGCGGATGGTGTAAATTCATGACACTGGGGCAACACCAAGTGGCATC 1586
Db      1203 GAAGCTGGTGCATTTGGCTTTAAATCCAGAACTTGGGGCACACTCTCTTCTGCAATC 1144
Qy      1587 GATCACTGCTTGACGCTGGCAGATGAATAGATGTGCAAGTTTGTATCCACACCGATACA 1646
Db      1143 AATCATGCGTTAGATGTTGCAGACAAATACGATGTGCAAGTCGCTATCCACACAGACT 1084
Qy      1647 GTCAATGAGCAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGCGCCATCCAT 1706
Db      1083 TTGAATGAAGCCGTTGCGTGAAGACACTATGGCAGCTATTGCGGACGCACTATGCAC 1024
Qy      1707 GCCTACCACTTGAAGGCGGGTGGAGGACACTCACTGATGTTATCACCATGGCAGGC 1766
Db      1023 ACTTTCCACACTGAAGGTGCTGGCGGGGACACGCTCTGATATATTAAAGTAGCTGGT 964
Qy      1767 GAGCTCAATATCTACCCCTCTCCACACCCCACTATTCCTATACATTTAATACGGTT 1826
Db      963 GAACACAACTCTCCCGCTTCCACTAACCCCACTATCCCTTCTCATGTGAATACAGAA 904
Qy      1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCTCCGCGAGGAT 1886
Db      903 GCAGAACACATGGACATGCTTATGTTGTGCCACCACTTGGATAAAGCAATTAAGAAGAT 844
Qy      1887 TTAACAATTTCTAAAGCCGTAATCCGCCCGGCTCTATCGCGGCTGAAGATGTCTCCAT 1946
Db      843 GTTCAGTTCGCTGATTCAAGGATCCGCCCTCAAAACCATTTGCGGTGAAGACACTTTGCA 784
Qy      1947 GATATGGTGTGATCGCGATGACAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGGCGAA 2006
Db      783 GACATGGGATTTTCTCAATCACAGCTCTGACTCTCAAGCTATGGGTGCTGTGGGTGAA 724
Qy      2007 GTGATTTCTCGAACTTGGCAGACTGCGGATAAGAAATTAATAAGAAATTTGGTAAAGCTTCT 2066
Db      723 GTTATCACTAGAACTTGGCHAAACAGCTGACAAAAACAAAAAGAAATTTGGCCGCTTGAA 664
```

```
Qy      2067 GAAGATGSCAAAGATAACGATAAATTTCCGCATTAAGCGCTACATCTCCAAATACACTATC 2126
Db      663 GAAGAAAAGGCGATAACGACAACTTCAGATCAAAACGCTACTTGTCTAAATACACCAAT 604
Qy      2127 AACCCTGCTTTACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATC 2186
Db      603 AACCAGCGATCGCTCATGGGATTAGCGAGTATAGGTTCTGTAGAAGTGGCAAGTG 544
Qy      2187 GCCGACTTGGTGGTGGGAATCCTGCTTTTGGCGTAAAAACCCAAAAATCGTGATCAA 2246
Db      543 GCTGACTTGGTATGTGGAGTCCCGCATTTCTTTGGGTTAAACCCCAACATGATCATCAA 484
Qy      2247 GCGGATATGGTGGTCTTCTCTGAATGGCGAATCTAAGCGCTGTGTGCCCCACTCCCCAA 2306
Db      483 GCGGGTTCTATGCGTTGAGTCAATGGGTAGCGCAACGCTTCTATCCTACCCCAACAA 424
Qy      2307 CCGGTTTATACCGGAAATGTTGGGCTACGCGCAAGCGGAATTTGACACACGATC 2366
Db      423 CCAGTTTATACAGAGAAATGTTGCTCATCATGTGTAAGCCAAATACGATGCACAAATC 364
Qy      2367 ACTTTGTTTCCAAAGTCGCTATGAAATGGCGTAAAGAAAAAGCTGGGCTTAGAGCGC 2426
Db      363 ACTTTGTTCTCAAGCGCTTATGACAAAGCATTTAAAGAGAAATTAGGCTTGAAGA 304
Qy      2427 CAAGTTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db      303 CAAGTTTCCCGTAAATAAATTTGCAGAAACATCACTAAAAAAAGACATGCAATTTCAACGAC 244
Qy      2487 AAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTC 2546
Db      243 ACTACCGCTCACTTGAAGTCAATCTGAAACTTACCATGTGTTGTTGTTGTTGTTGTTGTT 184
Qy      2547 TGCACCTTAACCCACCTCGCAAGTCCTCTAGCCAGCGGTACACTTTCTTCTAGG 2604
Db      183 GTAACTTCTAAACCGCAATAAAGTGAGCTTTGGCGCAACTCTTTAGCATTTTCTAGG 126

RESULT 15
AR054309
LOCUS      AR054309          2735 bp          DNA          linear          PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5837240.
ACCESSION AR054309
VERSION    AR054309.1  GI:5979886
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 2735)
AUTHORS     Lee, C.K., Monath, T.P., Ackerman, S.K., Thomas, W.D., Soman, G.,
            Kleanthous, H., Weltzin, R.A., Pappo, J., Ermak, T., Guirakhoo, F.,
            Bhagat, H. and Suesman, I.
TITLE       Multimeric, recombinant urease vaccine
JOURNAL     Patent: US 5837240-A 17-NOV-1998;
FEATURES    Location/Qualifiers
             source          1..2735
             /organism="unknown"
             /mol_type="unassigned DNA"

Query Match      37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred No. 2 4e-204;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;

Qy      182 AAAATTTTAAACAAAGGAGTAATAGTGAAACTCACACCCAAAGACAGAGGCTTAAGCTCAACCAA 301
Db      78 ACACTTTAAAGATAGGAGAAATGAGATGAACTCACCCCAAAAGAGTTAGATAAGTTGATG 137
Qy      242 TTATATTATCGGGCGGAAGTGGCTAGAAAGCGCAAGCGAGGAGGCTTAAGCTCAACCAA 301
Db      138 CTCCTACTACGCTGGAGAAATGGCTAAAAAACGCAAAAGAAAAAGGCAATTAAGCTTAAC 197
Qy      302 CCCGAAGCCATTGCTTTACATTAGTCCCATATTATTATGGACGAAGCGCGCTTGAAAAAAA 361
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 23:31:00 ; Search time 10210 Seconds
(without alignments)
13211.279 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgrgattttccaractt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156643

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	327.4	11.4	958	8	DR637438	DR637438 EST102806
2	324.2	11.2	946	8	DR637093	DR637093 EST102771
3	310.2	10.8	1509	4	CNS0ABR8	BX816461 Arabidops
4	309	10.7	872	7	CO122599	GR_Eb04F
5	306.2	10.6	925	8	DR633422	DR633422 EST102404
6	302.2	10.5	748	8	CK676965	Yd29h04
7	299.2	10.4	889	8	DR637386	DR637386 EST102801
8	296.4	10.3	796	8	DR633939	DR633939 EST102456
9	292	10.1	849	8	DR635599	EST102622
10	279.6	9.7	766	7	CV195867	CGF100344
11	274.2	9.5	986	6	CF826586	EST703968
12	273	9.5	784	9	AZ935182	AZ935182 BJ_Ba000
13	262.6	9.1	990	7	CO027797	CO027797 EST806181
14	261.2	9.1	684	7	CN907928	CN907928 030109ABL
15	261	9.1	921	6	CF714848	CF714848 CCAF343TO
16	261	9.1	946	8	DN809412	DN809412 73954613
17	259.6	9.0	713	7	CK757406	atr02-4ms
18	258.2	9.0	689	8	DR440006	DR440006 EST149_09
19	255.2	8.9	781	8	DR916924	EST110846
20	254.8	8.8	863	6	CD375065	PTMM00226
21	254.4	8.8	945	8	DN808832	DN808832 76945303
22	249.4	8.7	757	5	BU027777	QHG7L07.y

ALIGNMENTS

RESULT 1
DR637438
LOCUS
DEFINITION DR637438 958 bp mRNA linear EST 11-JUL-2005
EST1028063 FvM Gibberella moniliformis cDNA clone FVMAW64, mRNA
sequence.
ACCESSION DR637438
VERSION DR637438.1 GI:70712272
KEYWORDS
SOURCE
ORGANISM
Gibberella moniliformis
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE
1 (bases 1 to 958)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Kendra, D.P., Town, C.D. and Whitelaw, C.A.
Xentrac, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
JOURNAL
COMMENT
Contact: Brown, D.W.
USDA/ARS/NCAUR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncaur.usda.gov
TIGR sequence name: FVMAW64TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.
Location/Qualifiers
1. 958
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAW64"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="vector: pBluescript II SK(+); XR; Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was

ORIGIN

directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."

Query Match 11.4%; Score 327.4; DB 8; Length 958;
Best Local Similarity 59.9%; Pred. No. 3.5e-70;
Matches 565; Conservative 0; Mismatches 376; Indels 2; Gaps 1;

QY 1209 GATGCGTAAGCCCTCATATGGTTCGTGGGTGGGACAGAGCACTAGCAGGGGAAGGT 1268
|||||
Db 2 GACGCGTAACAGAGGCGATGTCGTCGGAAGTCGACGGATGCTGATAGCAGGTGAAGGA 61
|||||
QY 1269 ATGATTATTACCGTGGGGGAATCGATTCACACACCCTCTCTTCTCCACAACTTC 1328
|||||
Db 62 AAGATTGTGACCGCGCGCTATCGACAGCATATTCATTTATCTGCGCTCAGCAAGTA 121
|||||
QY 1329 CCTACCGTCTAGCCAAATGCGGTACCAACATGTTGGAGGGGACACAGTCTGTAGAT 1388
|||||
Db 122 CCGAAGCTCTTGCATCTGGTGTAAACCACTGCTTGGCGGTGGTACCGGCCCAAGTGT 181
|||||
QY 1389 GGCAGGAATGCGACTACTATCACTCCGGGCAAAATGGAATTCGACCGCATGTTGCGCGCA 1448
|||||
Db 182 GGAACGAGCAACTACTTGTACGCTGTGTCTCATTCATCGGTCAATGTTGCAGGCG 241
|||||
QY 1449 GCAGAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAA 1508
|||||
Db 242 TCGCATGAGCTTCTTATCAATATGTTATTAATCTGCAAAAGGTAATGATAGTTCTCTCTGAG 301
|||||
QY 1509 CAACTGTAGACAGTAGAAGGGGCGATGTTGTTTAAATGATGATGAGTCTGAGGCT 1568
|||||
Db 302 GGTCTGCGCGATGAGTCAATGCTGGTGTGTTGGGCTCAAGCTTCATGAGGACTGGGT 361
|||||
QY 1569 ACAACACCAAGTGCAGTCGATCACTCTTGGAGGTGGCAGATGAATACGATGTGCAAGTT 1628
|||||
Db 362 TGCATCTCTGCTGCTATTGACGCTTGTCTCAGTGTCTGTGATGAATTCGATTCATGAT 421
|||||
QY 1629 TGTATCCACACCGATACAGTCAATGAGGAGGAGTTATGTAGATGACACCCCTAAATGCAATG 1688
|||||
Db 422 CTATTTCACACTGACAGCTTAACGAGTCTGGCTTGTGCGAATCTACGATCGTCTTC 481
|||||
QY 1689 AACGGGCGGCATCCATGCTACCACTTACGAGGCGGTGGAGGACACTCACTGAT 1748
|||||
Db 482 AAGAAGCCGACAAATTCATCTTATCAACAGAGGGTGCAGGAGGTGGCCATGCTCCGGAT 541
|||||
QY 1749 GTTATCACCATTGCGAGCGAGTCAATATTTACCTCTCTCCACACCCCTCACTATTC 1808
|||||
Db 542 ATCACTCTCGTGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCTCAAGACCA 601
|||||
QY 1809 TATACCATTAATACGGTTGAGAACACTTAGACATGCTCATGACATGCCACCACTAGAC 1868
|||||
Db 602 TTCACAGCAATACTCTCGATGAGCATCTCGATATGCTTATGCTGCTGCCATCACTGTCC 661
|||||
QY 1869 AAACGATCCGGAGAGATTACAAATTTCTCAAAGCGTATCCGCCCGGCTCTATCGCG 1928
|||||
Db 662 AAGAATATCCAGAGGATGAGCTTTCGGCGAGAGCGGTATTCGTGCTGAAACCAATGCT 721
|||||
QY 1929 GCTGAGATGTGCTCCATGATATGGGTGTGATCGCATGACAGCTCGGATTCGCAAGCA 1988
|||||
Db 722 GCTGAGGATGTATTAACGCAAAAGCGCTATCAGCATGATGAGTCTGACTCGCAGGCT 781
|||||
QY 1989 ATGGGCGGTGCGAGCGCAAGTGTATTCCTGCAACTTTGCGAGATCGCGATGAAGATAAAAA 2048
|||||
Db 782 ATGGGCGGTGCGAGAGGTGCTTTTAAAGAACATGAATACTCGCATATAAATAGGTG 841
|||||
QY 2049 GAATTTGGTAAGTCTCTGAAGATG--GCAAGATAAAGGATAATTTCCGATTAAGCGCT 2106
|||||
Db 842 CAGAGGGTGTGTTGCGGAGGATGAAGGCACAGGGGCTGATTAATGCGCGGTGAACGCT 901
|||||
QY 2107 ACATCTCCAAATACATATCAACCCCGCTTTGACCCACGCGT 2149
|||||
Db 902 ATGTCAGCAAGTACTATTAAACCCAGCTATTGCTCAGGCGCT 944
|||||

RESULT 2

DR637093

LOCUS

DEFINITION

EST1027718 FvM Gibberella moniliformis cDNA clone FvMAT95, mRNA

sequence.

DR637093 946 bp mRNA linear EST 11-JUL-2005

DR637093.1 GI:70711927

ACCESSION

DR637093

VERSION

DR637093.1

KEYWORDS

EST.

SOURCE

Gibberella moniliformis

ORGANISM

Gibberella moniliformis

REFERENCE

1 (bases 1 to 946)

1815 N. University St, Peoria, IL 61604, USA

USDA/ARS/NCAUR

1815 N. University St, Peoria, IL 61604, USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brownw@ncaur.usda.gov

TIGR sequence name: FvMAT95TH

Seq primer: AAT TAA CCC TCA CTA AAG GG.

LOCATION/Qualifiers

1. 946

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="m3125"

/db_xref="taxon:117187"

/clone="FvMAT95"

/tissue_type="mycelia"

/clone_lib="FvM"

/notes="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour culture of Gibberella moniliformis grown on a 48-hour and a 72 hour, liquid GYM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match 11.2%; Score 324.2; DB 8; Length 946;
Best Local Similarity 59.7%; Pred. No. 2.2e-69;
Matches 563; Conservative 0; Mismatches 378; Indels 2; Gaps 1;

QY 1209 GATGCGTAAGCCCTCATATGGTTCGTGGGTGGGACAGAGCACTAGCAGGGGAAGGT 1268
|||||
Db 2 GACGCGTAACAGAGGCGATGTCGTCGGAAGTCGACGGATGCTGATAGCAGGTGAAGGA 61
|||||
QY 1269 ATGATTATTACCGTGGGGGAATCGATTTCACACCCACTTCTCTTCTCCACAACTTC 1328
|||||
Db 62 AAGATTGTGACCGCGGCGCTATCGACAGCATATTCATTTATCTGCGCTCAGCAAGTA 121
|||||
QY 1329 CCTACCGCTCTAGCCAATGGCTTACCAACCATGTTTGGAGCGGCACAGGTCTCTGTAGAT 1388
|||||
Db 122 CCGAAGCTCTTGATCTGTGTTAAACCAACCATGCTTGGCGGTGGTACCGGCCCAAGTGT 181
|||||
QY 1389 GGCACGAATGCGACTACTATCACTCCGGGCAAAATGGAATTCGACCGCATGTTGCGCGCA 1448
|||||
Db 182 GGAACGAGCAACTACTTGTACGCTGTCTCATTTACATGCGTCAATGTTGCAGGCG 241
|||||
QY 1449 GCAGAAGAGTATTCTATGAATGTGGGCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAA 1508
|||||
Db 242 TCGCATCAGCTTCTCAATATTTGGTATTACTGCAAAAGGTAATGATAGTTCTCTCTGAG 301
|||||

```
Qy 1509 CAACTTGTAGAACAGTAGAGCGGGCGAATTGGTTTAAATTCATGATGAAGACTGGGGC 1568
Db 302 GGTTCGCGGATCAGGTCAATGCTGGTGTGTGGCTCAAGCTTCATGAGGACTGGGGT 361
Qy 1569 ACAACACCAAGTGCATCGATCACTGCTTGAGGGTGGCAGATGAATACGATGTGCAAGTT 1628
Db 362 TGCACATCTCTGCTGCTATGACGGTTGTCTCAGTGTCTGTGATGAATTCGATATTCATGT 421
Qy 1629 TGTATCCACACCGCATACAGTCAATGAGGAGGTTATGTAGATGACACCCCTAAATGCAATG 1688
Db 422 CTTATTCACACTGACACGCTTACGAGTCTGGCTTTGTGCAATACGATCGTCTTC 481
Qy 1689 AACGGGCGGCCCATCCATGCTTACCAATGAGGAGCGGGTGGAGGACACTCACCTGAT 1748
Db 482 AAGAACCGCACAAATTCATCTATCACAGAGGGTGCAGGAGGTGCCATGTCTCCGAT 541
Qy 1749 GTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACCCACCCCACTATTC 1808
Db 542 ATCATCTCGGTGTAGAGCATCAAAATGTTCTGCCATCATCGACCAACCCCTACAAGACCA 601
Qy 1809 TATACCAATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGAC 1868
Db 602 TTCACAGGCAATCTCTCGATGAGCATCTCGATATGCTTATGCTGCTGCCATCACTGTCC 661
Qy 1869 AAACGATCCGCGAGGATTTACAATTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCG 1928
Db 662 AAGAATATCCAGAGGATGTAGCTTTCGGGAGAGCGGTATTCGTGCTGAACCATTTGCT 721
Qy 1929 GCTGAAGATGTGCTCCATGATATGGGTGATCGGATGACAGCTCGGATTCGCAAGCA 1988
Db 722 GCTGAGGATGTATTACAGCAAAAGCGCTATCAGCATGATGAGCTCTGACTCGCAGGCT 781
Qy 1989 ATGGGCGGTGCGAGCAAGTATTCCTCGAACTTGCAGACTCGGATAGATAAAAA 2048
Db 782 ATGGGCGGTGCGAGAGGTGCTTTAAGAACTAGGAATCTGCGCATAAAAAATAGTG 841
Qy 2049 GAATTTGGTAAAGCTTCTCTGAAGAT--GGCAAAGATAACGATAATTTCCGCAATTAAGCGCT 2106
Db 842 CAGAGGGTTGGTTGCGGAGGATGAGGCACAGGGGCTCATATGCGCGTGTAAAGCT 901
Qy 2107 ACATCTCCAAATACATATCAACCCCGCTTTGACCCACCGCGT 2149
Db 902 ATGTCAGCAAGTATACTATTACCCAGCTATTGCTCAGGGCTT 944

RESULT 3
CNSOABR8 1509 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTPGH4ZD08 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX816461
VERSION BX816461.1 GI:42471655
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1509)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
```

```
COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
Location/Qualifiers
source 1..1509
/mol_type="mRNA"
/db_xref=taxon:3702"
/clone="GSLTPGH4ZD08"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
gene 1..1509
/gene="Atlg67550"
ORIGIN
Query Match 10.8%; Score 310.2; DB 4; Length 1509;
Best Local Similarity 59.1%; Pred. No. 7.7e-66;
Matches 602; Conservative 0; Mismatches 413; Indels 4; Gaps 4;
Qy 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGCATACA 1646
Db 1 GACAATTTGTTTGGCAGTTGCAGAAATATGATATCCAGGTGAACATTCATATCTGACACC 60
Qy 1647 GTCAATGAGGCGAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGGCCATCCAT 1706
Db 61 TTAATGAATCCGGTTTGTGTGACACACTATCAATGCAATTCGTGGGAGAACATATCAT 120
Qy 1707 GCCTACCACATTTGAGGGAGCGGGTGGAGCACACTCACCTGATGTTTATCACCATGGCAGGC 1766
Db 121 ACATATCACAGTAAGGTGCTGGTGGATGATGCACAGATATCATTAGAGTTTGTGGA 180
Qy 1767 GAGCTCAATATTTACCCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAATACGGTT 1826
Db 181 GTGAAAATGTACTCCCGTCTATCAACCAACCCCACTATATACGAAAATACTGTA 240
Qy 1827 GCAGAACACTTTAGACATGCTCATGACATGCGACCACTAGACAAAGCATTCGCCGAGGAT 1886
Db 241 GATGAACATCTTGACATGCTGATGTTTGGCCATCACCTTGACAAGAACATCCCAGAAGAT 300
Qy 1887 TTACAATTTTCTCAAGCCGTATCGCCCGGCTCTATCGCGCTGA-AGATGTGCTCCA 1945
Db 301 GTAGCTTTTGTGTAATCAAGAAATAGAGCTGAAACAATAGCTGCGAGAGAGATATATGCA 360
Qy 1946 TGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGCGAGCGGA 2005
Db 361 TGATATGGAGCAATCAGCATATCTCTCCGACTCAAGCAATGGGTGCGATTGGAGA 420
Qy 2006 AGTGATTTCTCGAACTTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTC 2065
Db 421 GGTAATAAGCAGAACTTTGGCAGACAGCTGCAAGATGAAGGTCTCAGAGGGGAGCCATTAA 480
Qy 2066 TGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGC-GCTACATCTCAATACACTA 2124
Db 481 TCCTAAACATGCTGATGATGATAACTCAAGGATCAAGCAGATACATTTGCAAGATGACACTA 540
Qy 2125 TCACCCCGCTTTTGACCCACCGCTGAGCGAGTATATCGGCTCTGTGGAGAGGGGCAAGA 2184
Db 541 TAAACCCAGCTATAGCCCAATGGGTTTGAGACTTAATTTGGCTCGGTGAGGTGAAGAGC 600
Qy 2185 TCGCCGACTTGGTGGTGGTAATCTCGCTTTT-TTGGCGTAAAAACCCCAAAATCGTGATC 2243
Db 601 TGGCTGATCTTGTTATATGGCAGCGCGCTTTCTTCGGAGCAAAACCCAGAAATGATAATC 660
```

```
FEATURES
Location/Qualifiers
source 1..1509
/mol_type="mRNA"
/db_xref=taxon:3702"
/clone="GSLTPGH4ZD08"
/tissue_type="Hormone Treated Callus"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
gene 1..1509
/gene="Atlg67550"
```


REFERENCE 1 (bases 1 to 925)
 BROWN, D.W., CHEUNG, F., PROCTOR, R.H., BUTCHKO, A.E., ZHENG, L., LEE, Y.,
 UTTERBACK, T., SMITH, S., FELDBLYUM, T., GLENN, A.E., PLATTNER, R.D.,
 KENDRA, D.F., TOWN, C.D. and WHITELAW, C.A.
 Analysis of 87,000 expressed sequence tags reveals alternatively
 spliced introns in multiple genes of the fumonisin gene cluster
 Unpublished (2005)
 Contact: Brown, D.W.
 USDA/ARS/NCAR

1815 N. University St, Peoria, IL 61604, USA
 Tel: 309 681 6230
 Fax: 309 681 6689
 Email: brownw@ncar.usda.gov
 TIGR sequence name: FVMA534TH
 Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES
 source
 1..925
 /organism="Gibberella moniliformis"
 /mol_type="mRNA"
 /strain="m3125"
 /db_xref="taxon:117187"
 /clone="FVMA534"
 /tissue_type="mycelia"
 /clone_lib="FvM"
 /notes="Vector: pBlueScript II SK(+); XR; Site 1: EcoRI;
 Site 2: XhoI; anamorph: Fusarium verticillioides. Library
 FvM was prepared from pooled RNA obtained from a 48-hour
 and a 72 hour, liquid GRAM culture from strain M-3125.
 Cultures were vacuum filtered and the mycelial mats were
 frozen in liquid nitrogen, ground to a powder, and then
 added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
 approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
 directionally ligated into the pBlueScript II SK(+) XR
 vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN
 Query Match 10.6%; Score 306.2; DB 8; Length 925;
 Best Local Similarity 59.9%; Pred. No. 6.7e-65;
 Matches 548; Conservative 0; Mismatches 363; Indels 4; Gaps 2;

1312 TTCTCCACAAATCCCTACCGCTCTAGCCATGGCGTTACCAACCATGTTGGAGCG 1371
 DB 10 TCTGCCCTCAGCAAGTACCGAAGCTCTTGCACTGGTGTAAACCAACATCTTGGCGTG 69
 QY 1372 GCACAGGTCTGTAGATGGCAAGATGCGATCTACTATCATCTCCGGGCAATGGAATTCG 1431
 DB 70 GTACCGGCCCAAGTGTGGAACGAAACGCAACTACTTGTACGCTGTGTCTCATTCATGC 129
 QY 1432 ACCGATGTTCGGCGCAGCAGAGAGTATCTATGATGTGGCTTTTGGGCAAGGCA 1491
 DB 130 GTCAATGTTCGAGCGGTGCGATCAGCTTCCATCAATATGTTATTAAGGCAAGGTA 189
 QY 1492 ATAGCTCTAGCAAAACAACTGTGTAGAACAGTAGAGCGGCGGATGTTTAAAT 1551
 DB 190 ATGATAGTCTCTGAGGTCTCGCGATCAGTCAATGCTGTGTGTGGCTCAAGC 249
 QY 1552 TGCATGAAGATCGGGCACAACACCAAGTCGATCGATCATCTGCTGAGCGTGGCAGATG 1611
 DB 250 TTCATGAGGACTGGGTGTCACCTCTGCTCTATTGACGCTTGTCTCAGTGTCTGTATG 309
 QY 1612 AATACATGTGGAAGTTTGTATCCACCGATCAGTCAATGAGGCGAGTTATGATG 1671
 DB 310 AATTCGATATTCATGCTCTTATTCACACTGACACGCTTAAACGAGTCTGGCTTTGTCGAAT 369
 QY 1672 ACACCTTAATGAATGAACGGGCGCCATCATCTGCTACCATGTTGAGGCGGCGTG 1731
 DB 370 CTACGATTCGCTTTTCAAGAACCGCAAAATTCATCTTATCACAGAGGGTGCAGAG 429
 QY 1732 GAGGACACTCACCTGATGTTATCACCATGGCAGCGAGCTCAATATTTCTACCTCTCTCCA 1791
 DB 430 GTGGCCATGCTCGGATATCACTTCGTGTAGAGCATCAAAATGTTCTGCCATCATCGA 489

QY 1792 CCACCCCACTATTCCCTATACCATTAATACGGTTGCAGAACACTTAGACATGCTCATGA 1851
 DB 490 CCAACCTTACAGAACCATTCACAGCAATATCTCGATGAGCATCTCGATATGCTTATGG 549
 QY 1852 CATGCCACCTAGACAAACGCAATCCGAGGATTTTCAATTTTCTCAAAGCCGTATCC 1911
 DB 550 TCTGCCATCATCTGTCCAAGATATCCAGAGGATGTAGCCTTCGGGAGAGCCGTATTC 609
 QY 1912 GCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGGGTGTGTCGGATGACAA 1971
 DB 610 GTGCTGAACCATTCGCTGAGGATGTATACACGACAAAGCGCTATCAGCATGATGA 669
 QY 1972 GCTCGATTCCGAAGCAATGGGGCTGCAGGGAAGTATTCCTCGAACTTGGCAGACTG 2031
 DB 670 GCTCTGACTCGAGGCTATGGGCGCTTCGAGAGAGTGTGTTTAAAGAACATGGAATAC 729
 QY 2032 CGGATAGAATAAAAGAAATTTGTAAGCTTCCTGAAGAT--CGCAAGATAACGATA 2088
 DB 730 CGCATAAAATAAGGTGCAGAGGGGTGTTGCCGAGGATGAGGCAACAGGGCTGATA 789
 QY 2089 ATTTCCGATTAAAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACCCACGGG 2148
 DB 790 ATGCGCGTGTAAAGCTATGTCAAGATATATCTATTAAACCCAGCTATGCTCAGGGCT 849
 QY 2149 TGACCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCGA-CTTGGTGTGTGGAAT 2207
 DB 850 TTGGACATGTCTATTTGGAGCATTCAGGTGTGAAGTTTGTCTGATCTTCTTTTGGGAT 909
 QY 2208 CTTGCCCTTTTGGC 2222
 DB 910 CTTGCCGTGTTGGC 924

RESULT 6
 CX676965 748 bp mRNA linear EST 19-JAN-2005
 LOCUS ydd29h04.y1 Sea urchin EST Lib1 Strongylocentrotus purpuratus cDNA
 clone ydd29h04 5' similar to TR:Q9SR52 Q9SR52 PUTATIVE UREASE. ;
 mRNA sequence.

ACCESSION CX676965.1 GI:57936920
 VERSION EST.
 KEYWORDS Strongylocentrotus purpuratus
 SOURCE Strongylocentrotus purpuratus
 ORGANISM Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoidea; Euechinoidea; Echinacea; Echinoida;
 Strongylocentrotidae; Strongylocentrotus.
 1 (bases 1 to 748)
 Coffman, J.A., Robertson, A.J., Clifton, S., Pape, D., Hillier, L.,
 Martin, J., Wylie, T., Dante, M., Meyer, R., Theising, B., Bowers, Y.,
 Gibbons, M., Ronko, I., Tsagaris, V., Ritter, E., Kennedy, S. and
 Wilson, R.
 WashU Sea Urchin EST Project
 Unpublished (2004)
 Contact: Dr. James A. Coffman
 WashU Sea Urchin EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu

DNA sequencing by: Washington University Genome Sequencing Center
 Seq primer: -28RPPOT
 High quality sequence stop: 680.
 Location/Qualifiers
 source
 1..748
 /organism="Strongylocentrotus purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="ydd29h04"
 /lab_host="DH10B"
 /clone_lib="Sea urchin EST Lib1"
 /note="Vector: pCMVSPORT6.1 (Invitrogen); Site_1: NotI;

Site_2: Smal; Arrayed normalized library of full-length cDNAs representing blastula stage transcriptome of the sea urchin Strongylocentrotus purpuratus, cloned into the vector pCMVSPORT6.1 (Invitrogen)"

ORIGIN	Query Match	Score	DB 8;	Length	748;
	Best Local Similarity	63.1%;	Pred. No. 6.3e-84;		
	Matches	466;	Conservative	0;	Mismatches 273; Indels 0; Gaps 0;
QY	1648	TCAATGAGCGAGTATGTAGATGACACCTTAATCAATGAACGCGCGGCCATCCATG	1707		
Db	9	TCATATGAATCTCTGCTGTGTGGAAGATACCATAGCTGCTTTGAGGCGCGCACCATCCACA	68		
QY	1708	CCTACACATTCAGGAGCGGTGGAGGACATCTCACCTGATGTTATCACCATGCGAGGCG	1767		
Db	69	CTATCACTCGAGGAGTGCAGGAGGAGCCACGCCCTTGACATCATGAAGTAGTGGCG	128		
QY	1768	AGCTCAATATCTACCTCTCTCACACCGCCCACTATTCCTATACATTAATACGGTTG	1827		
Db	129	TGCCCAATGTGTACCTCTGCTTACCAACCCCAACCGCCCTTCACAGTGAACACCATCG	188		
QY	1828	CAGAACCTTACATGCTCATGACATGCGACCATGCGACATGACACACGATCCGGAGATT	1887		
Db	189	ATGAACACCTGGATATGCTCATGGTCTGCCACCATCTTGACAAGAAATCTAAAGGAAGCG	248		
QY	1888	TACAAATTTCTCAAGCCGTATCCGCCCGGCTCTATCGGGCTGAAGATGTCTCCATG	1947		
Db	249	TGCGCTTTGCAGAGTCTCGCATCAGAGCCGAACCAATTCGTCTGAAGATATCTTCATG	308		
QY	1948	ATATGGGTGTATCGCGATGACAAAGTCGGATTCGGAACAAATGGGGCGTGCAGGCGAAG	2007		
Db	309	ATCTGGGAGCAATCAGTATTTGGCTCCGATTCGCAGGCCATGGGGCGTGTGCGTGAGG	368		
QY	2008	TGATCTCTCGAATTCGGAGCTGCGGATGAAGATATAAAGAAATTTGGTAAGCTTCCTG	2067		
Db	369	TGATCACCCTGATACATGGCAGCGGCAGATAAGATGAAGATTTTTCAGAGGCGAGCTGTCA	428		
QY	2068	AAGATGGCAAGATACAGTAATTTCCGCATTAAAGCGCTACATCTCCAAATACACTATCA	2127		
Db	429	AGGAACGGGTGACACGACCACTTAAGGTCGAGGTATACATGTCGAAGTACACGGTCA	488		
QY	2128	ACCCGCTTTGACCCACCGCGTGAAGGATATATCGGCTCTGTGGAAGAGGCGCAAGATCG	2187		
Db	489	ACCCGCGCAATGCCCCACGCGATGGGCCATCTCATTTGGCTCTGTACAGTTGGTAAATGG	548		
QY	2188	CGACTTGTGTGTGGAATCTCGCTTTTGGCGTAAACCCAAATTCGTGATCAAAG	2247		
Db	549	CTGACCTGGTCTCTGGAATCCGGCATCTTTCGGAGCTAAACCTGACCTCATCAAAG	608		
QY	2248	GGGTATGTGTCTTCTCTGAAATGGCGATTTCTAACGGCTGTGCGCCACTCCCAAC	2307		
Db	609	GAGGTTACATGCTTGGGCACAAATGGGCATGCCCCAATGATCCATCCCGACCCCTGAAC	668		
QY	2308	CGGTTATTACCGCGAAATGTTTGGGCATCACCGGCAAGGCGAAATTTGACACCAAGATCA	2367		
Db	669	CTGTCAAGACACGAGAAATGTTTGGTCTTACGGGAAGTCCATCGGTGAGAACTCTGTCA	728		
QY	2368	CTTTTGTTCCTCAAGTCGC	2386		
Db	729	TCTTTGTGTCCTCAAGCAGC	747		

RESULT 7
DR637386
LOCUS
DEFINITION
EST1028011 FvM Gibberella moniliformis cDNA clone FVMAV95, mRNA
sequence.
ACCESSION
DR637386
VERSION
DR637386.1 GI:70712220
KEYWORDS
EST.
SOURCE
Gibberella moniliformis
ORGANISM
Gibberella moniliformis

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
1 (bases 1 to 889)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D., and Whitelaw, C.A.
Analysis of 87,000 expressed sequence tags reveals alternatively
spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAUR
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brownw@ncaur.usda.gov
TIGR sequence name: FVMAV95TH
Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES
source
1..889
Location/Qualifiers
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAV95"
/tissue_type="mycelia"
/clone_lib="FvM"
/notes="Vector: pBlueScript II SK(+); XR: Site 1: EcoRI;
Site_2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to TRIzol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml TRIzol. The cDNA was
directionally ligated into the pBlueScript II SK(+)-XR
vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN	Query Match	Score	DB 8;	Length	889;
	Best Local Similarity	59.5%;	Pred. No. 3.7e-63;		
	Matches	524;	Conservative	0;	Mismatches 353; Indels 3; Gaps 1;
QY	1312	TTTCTCCACAACTTCCCTACCGCTCTAGCAATGGCTTACCAACCATGTTTGAGGGG	1371		
Db	10	CTGCCCCGACAGTACCCGAGCTCTTGCACTGTGTGTAAACCACTGTTGCGGTG	69		
QY	1372	GCACAGTCTCTGTAGATGGCAAGTACGACTACTATCACTCCGGGCAAAATGGAATTC	1431		
Db	70	GTACCGGCCCAAGTGTGGAAAGCAAGCAACTACTTGTACGCTGTGCTCATTAATGC	129		
QY	1432	ACCGCATGTTGCGCGCAGCAGAGAGTATTCATGAATGTGGCTTTTGGGCAAGGCA	1491		
Db	130	GTCAAAATTTGCGAGCGCTGCGATCAGCTTCCATCAATATTGGTATTACTGCAAGGTA	189		
QY	1492	ATAGCTCTAGCAAAACAACTTGTAGAACAGTAGAGCGGCGGCGATTGGTTTAAAT	1551		
Db	190	ATGATAGTCTCTCGAGGGTCTGCGCGATCAGGTCATGCTGGTGTGTCCTCAAGC	249		
QY	1552	TGCATGAAGACTGGGGCACAACACCAAGTGCATCGATCACTGTGTAGCGGTGGCAGATG	1611		
Db	250	TTCATGAGGACTGGGGTTGCACCTCTGCTGCTATTGACGCTTGCTCAGTGTCTGTGATG	309		
QY	1612	AATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGCGAGGTATGATG	1671		
Db	310	AATTCGATATTCAATGTCTTATTACACATGACCGCTTAACGAGTCTGGCTTTGTGCAAT	369		
QY	1672	ACACCTTAATGCAATGAACGGGCGCCATCCATGCTTACCACTTTCAGGAGCGGTG	1731		
Db	370	CTAGCATCGCTGCTTTTCAAGAACCGCAATTCATCTATCACACAGGGGTGCGAGG	429		
QY	1732	GAGGACACTCACCTGATGTTTATCAACATGGGCGGAGCTCAATATTCTACCTCTCTCA	1791		

Db 430 GTGGCAGATCTCCGGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCCATCATCGA 489

Qy 1792 CCACCCCCACTATTCCTATACATTAATACGGTTTCAGAACACTTTAGACATCTCATGA 1851

Db 490 CCAACCTTACAGACCATTCACACGCAATACTCTCGATGAGCATCTCGATATGCTTATGG 549

Qy 1852 CATGCCACCACTAGACAAACGATCCGCGAGATTTACAAATTTCTCAAAGCGGTATCC 1911

Db 550 TCTGCCATCACTTGTCCAAGATATCCAGAGGATGTAGCTTCGCGGAGAGCGGTATTC 609

Qy 1912 GCCCGGCTCTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCATGACAA 1971

Db 610 GTGCTGAACCAATCTGCTGAGGATGTATTACACGCAAGGCGCTATCAGCATGATGA 669

Qy 1972 GTCGGATTTCGAAGCAATGGGCGGTGACGGGAGTGTCTCTGAACTTTGGCAGACTG 2031

Db 670 GCTCTGACTCGAGGCTATGGGCGGTGCGGAGAGTGTCTTTTAAAGCAATGAATACTG 729

Qy 2032 CGGATAAAGATAAAAAAGATTTGGTAACTTCTCTGAAGT---GGCAAGATGAACGATA 2088

Db 730 CGCATAAAAAATAGGTGACAGGGTGTGTGCGGAGGATGAGGSCACAGGGGCTGATA 789

Qy 2089 ATTTCGGCAATTAAGCGCTACATCTCAAAATACATATCAACCCGCTTTGACCCAGCGG 2148

Db 790 ATGCGCGTGTAAAAAGCTATGTGACGCAAGTATATTAACCCAGCTATTGCTCAGGGCT 849

Qy 2149 TGAGCGAGTATATCGCTCTGTGGAAGAGGCGCAAGATCGC 2188

Db 850 TTGGACATGTCATTTGAAGCATTTGAGGTGGAAAGTTTGC 889

RESULT 8

LOCUS DR633939

DEFINITION EST1024564 FvM Gibberella moniliformis cdna clone FvMA846, mRNA

ACCESSION DR633939

VERSION DR633939.1 GI:70708773

KEYWORDS EST.

SOURCE Gibberella moniliformis

ORGANISM Gibberella moniliformis

REFERENCE 1. (bases 1 to 796) Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y., Utterback, F., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D., Kendra, D.F., Town, C.D. and Whitelaw, C.A.

AUTHORS Analysis of 87,000 expressed sequence tags reveals alternatively spliced introns in multiple genes of the fumonisin gene cluster

TITLE Unpublished (2005)

JOURNAL Contact: Brown, D.W.

COMMENT USDA/ARS/NCAUR

USDA 1815 N. University St, Peoria, IL 61604, USA

Tel: 309 681 6230

Fax: 309 681 6689

Email: brown@ncaur.usda.gov

TIGR sequence name: FvMA846TH

Seq primer: AAT TAA CCC TCA CTA AAG GG.

FEATURES

source

1..796

/organism="Gibberella moniliformis"

/mol_type="mRNA"

/strain="M3125"

/db_xref="taxon:117187"

/clone="FvMA846"

/tissue_type="mycelia"

/clone_lib="FvM"

/note="Vector: pBlueScript II SK(+) XR; Site 1: EcoRI; Site 2: XhoI; anamorph: Fusarium verticillioides. Library FvM was prepared from pooled RNA obtained from a 48-hour and a 72 hour, liquid GYAM culture from strain M-3125. Cultures were vacuum filtered and the mycelial mats were

frozen in liquid nitrogen, ground to a powder, and then added to TRIzol Reagent (Invitrogen, Carlsbad CA) at approximately 1 g mycelia per 10 ml TRIzol. The cDNA was directionally ligated into the pBlueScript II SK(+) XR vector (cDNA Synthesis Kit; Stratagene)."

ORIGIN

Query Match 10.3%; Score 296.4; DB 8; Length 796;

Best Local Similarity 60.8%; Pred. No. 1.8e-62;

Matches 483; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

Qy 1243 GCACAGACACTAGCAGCGGAAGGTATGATTATACGGCTGGGGGAATCGATTACACACA 1302

Db 3 GCACGGATGTCGTAGCAGGTGAAGAAAGATTGACCGGGGCGCTATCAGCAGCGATA 62

Qy 1303 CCCACTTTCCTTCTCCACAACAATTCCTACCGCTCTAGCCAAATGGCGTTTACAACCATGT 1362

Db 63 TTCACITTTATCTGCCCTCAGCAAGTACCGAAGCTCTTGCAATCTGGTGTAAACCATGTC 122

Qy 1363 TTGGAGCGGCACAGGTCTCTGTAGATGGCAGCAATGCGACTTACTATCATCTCCGGCAAT 1422

Db 123 TTGGCGGTGTTACCGGCCCAAGTGTCTGAAAGCAAGCAACTTACTTGTACGCTTGGTGCTC 182

Qy 1423 GGAACCTTGACCGCATGTTTGGCGGCACAGAGAGTATTCTATGAATGGGCTTTTGG 1482

Db 183 ATTACATGCGTCAAAATGTTGCAGCGCTGCGATCAGTTCCTTATCAATATTGGTATTACTG 242

Qy 1483 GCAAGGCGCATAGCTCTAGCAAAAAACAATCTGTGAGAACAGTAGACGCGCGCGATTG 1542

Db 243 GCAAGGTAAATGATAGTTCCTCGAGGTCTGCCGATCAGGTCAATGCTGGTCTGTG 302

Qy 1543 GTTTTAAATTCATGAAGACTTGGGGCACAAACCAAGTGGCATCGATCACCTGCTTGAGCG 1602

Db 303 GCCTCAAGCTTCATGAGGACTTGGGTTGCACTCTCTGCTGCTATTGACGCTTGTCTCAGTG 362

Qy 1603 TGCAGATGAATACGATGTGCAAGTTTGTATCCACACGATACAGTCAATGAGGAGGTT 1662

Db 363 TCTGTGATGAATTCGATATTCAATGTCTTATTCACACTGACACGCTTAAACGAGTCTGGCT 422

Qy 1663 ATGTAGATGACACCTAAATGCAATGAACGGGCGGCCATCCATCCCTACCACATTGAGG 1722

Db 423 TTGTCGAATCTACGATCGCTGCTTTCAAGAACCCGCAAAATTCATACTTATCACACAGG 482

Qy 1723 GACGGGTGAGGACACTCACCTGATGTATTCACATGCGAGGCGAGCTCAATATTCTAC 1782

Db 483 GTGCAGAGGTGGCATGCTCCGGATATCATCTCCGTGGTAGAGCATCAAAATGTTCTGCG 542

Qy 1783 CTTCTCCACACCCCACTATTTCCTTATACCAATTAATACGGTTTCAGAAACACTTAGACA 1842

Db 543 CATCATCGACCAACCCCTACAAGACCAATTCACACGCAATACTCTCGATGAGCATCTCGATA 602

Qy 1843 TGCTCATGATGCGCACCTACAGCAAAAGCATCCGGAGGATTACAAATTTCTCAA 1902

Db 603 TGCTTATGCTGCGCATCACTTGTCCAAGAAATATCCAGAGAGGTATGACCTTCGCGGAGA 662

Qy 1903 GCGTATCCGCGCGCTCTATCGCGCTGAAGATGTCTCCATGATATGCTGTGATCG 1962

Db 663 GCGTATTCGTGTGTAACCAATGCTGCTGAGGATGTATTACACGCAAAAGCGGTATCA 722

Qy 1963 CGATGACAAGCTCGGATTTCGCAAGCAATGGGCGCTGAGGCGGAAAGTGAATTCCTCGAACTT 2022

Db 723 GCATGATGAGCTCTGACTCGCAGGCTATGGCGCTATGGCGAGAGGTGCTGTTTAAAGACAT 782

Qy 2023 GGCAGACTCGGAT 2036

Db 783 GGAATACTGCGCAT 796

RESULT 9

LOCUS DR635599

DEFINITION EST1026224 FvM Gibberella moniliformis cdna clone FvMAJ96, mRNA

sequence.

```
ACCESSION DR635599
VERSION GI:70710433
KEYWORDS EST.
SOURCE Gibberella moniliformis
ORGANISM Gibberella moniliformis
REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS Hypocromycetidae; Hypocreales; Nectriaceae; Gibberella.
TITLE 1 (bases 1 to 849)
Brown, D.W., Cheung, F., Proctor, R.H., Butchko, A.E., Zheng, L., Lee, Y.,
Utterback, T., Smith, S., Feldblyum, T., Glenn, A.E., Plattner, R.D.,
Kendra, D.F., Town, C.D. and Whitelaw, C.A.
JOURNAL Analysis of 87,000 expressed sequence tags reveals alternatively
COMMENT spliced introns in multiple genes of the fumonisin gene cluster
Unpublished (2005)
Contact: Brown, D.W.
USDA/ARS/NCAR
USDA
1815 N. University St, Peoria, IL 61604, USA
Tel: 309 681 6230
Fax: 309 681 6689
Email: brown@ncar.usda.gov
TIGR sequence name: FVMAJ98TH
Seq primer: AAT TAA CCC TCA AAG GG.
FEATURES
source
1..849
Location/Qualifiers
/organism="Gibberella moniliformis"
/mol_type="mRNA"
/strain="m3125"
/db_xref="taxon:117187"
/clone="FVMAJ96"
/tissue_type="mycelia"
/clone_lib="FvM"
/note="Vector: pBlueScript II SK(+); Site 1: EcoRI;
Site 2: XhoI; anamorph: Fusarium verticillioides. Library
FvM was prepared from pooled RNA obtained from a 48-hour
culture and a 72 hour, liquid GYM culture from strain M-3125.
Cultures were vacuum filtered and the mycelial mats were
frozen in liquid nitrogen, ground to a powder, and then
added to Trizol Reagent (Invitrogen, Carlsbad CA) at
approximately 1 g mycelia per 10 ml Trizol. The cDNA was
directionally ligated into the pBlueScript II SK(+) XR
vector (cDNA Synthesis Kit; Stratagene)."
```

```

Query Match 10.1%; Score 292; DB 8; Length 849;
Best Local Similarity 60.2%; Pred. No. 2.3e-61;
Matches 503; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

QY 1317 CCACAACTTCCCTACCGCTCTAGCAATGGCGTTACAAACCATGTTTGGAGCGGCACA 1376
DB 3 CTTCAAGATACCGAAGCTTTGGATCTGTCTTAACCAACCATGTTGGCGGTGTACC 62

QY 1377 GGTCTGTAGATGGCAAGATGGACTACTATCACTCCGGGCAAAATGGAACTTGCACCGC 1436
DB 63 GGCCCAAGTGCTGGAACGAACCACTACTTGTACGCTGTGCTCATTACATGGGTCAA 122

QY 1437 ATGTTGGCGGAGCAGAAAGATTTCTATGAATGTGGCTTTTGGGCAAAAGCAATAGC 1496
DB 123 ATGTTGGAGCGGTGGCATCAGCTTCCATCAATTAATTTGGTATTAATGCGCAAGGTAATGAT 182

QY 1497 TCTAGCAAAACAACCTTGTAGAACAGTAGAAGCGGCGCGATTGTTTTAAATTCAT 1556
DB 183 AGTTCTCTGAGGCTGCGCGATCAGGTCATCTGTTGCTGTGGCTCAAGCTTCAT 242

QY 1557 GAAGACTGGGGCACAACACCAAGTGGATCGATCGACTGCTTGAACGCTGGCGAGATGAATAC 1616
DB 243 GAGGACTGGGTTGCACTCTGCTGCTATTGACGCTGTCTCACTGCTGTGTGATGATTC 302

QY 1617 GATGTGCAAGTTGTATTCACACCGGATCAGTCAATAGGACGAGGTATGTAGATGACACC 1676
DB 303 GATATTCAATGTCTTATTCACACTGACAGCTTAAAGAGTCTGGCTTTGTGCAATCTACG 362

QY 1677 CTAATGCAATGNAACGGCGGCCATCCATGCCCTACCACATTCAGGGAGCGGTGGAGGA 1736
```

```

363 ATCGCTGCTTTTCAAGAACCGCACAAATTCATACTTATCACACAGAGGGTGCAGAGGTGGC 422
1737 CACTCACCTGATGTTATCACCATGGCGAGGAGCTCAATATCTTACCTCTCTCCACCAACC 1796
423 CATGCTCCGGATATCATCTCCGGTAGAGCATCAAAATGTTCTGCCATCATCGACCAAC 482
1797 CCACATATTCCCTATACCATTAATACGTTTGCAGAACACATTTAGACATGCTCATGACATGC 1856
483 CTTACAAGACCATTCACACCAATACTCTCGATGAGCATCTCGATATGCTTATGGTCTGC 542
1857 CACCACCTAGACAAACGATCCGCGAGGATTTCAATTTCTTCAAAGCGGTATCGCCCC 1916
543 CATCACTTGTCCAAAGATATCCAGAGGATGTAGCTTCCGCGAGAGCCGTATTGCTGCT 602
1917 GGCTCTATCGGGCTGAAGATGTCTCCATGATATGGTGTGATCGCGATGACAGCTCG 1976
603 GAAACCATTCTGCTGAGGATGATTACACGACAAAGCGCTATCAGCATATGAGCTCT 662
1977 GATTGCAAGCAATGGGGCGTCGAGCGAAGTGAATCTCGAACTTGGCAGACTGCGGAT 2036
663 GACTCGAGGCTATGGCCGTTGCGGAGAGTCTGTTTAAAGAACATGGAATCTGCGCAT 722
2037 AGAATAAAAAAGAAATTTGGTAAAGTTCCTGAAGAT---GGCAAAGATAACGATAATTC 2093
723 AAAATAAGGTGCAGAGGGGTTGGTTCCGCGAGGATGAGGCGACAGGGGCTGATAATCG 782
2094 CGCATTAAGCGCTACATCTCCAATACATCATCAACCCCGCTTTGACCCACGGCGT 2149
783 CGTGTAAACGCTATGTGACNAAGTATACTATTAAACCCAGCTATTGCTCAGGGCTT 838
```

```

RESULT 10
CV195867
LOCUS 766 bp mRNA linear EST 15-SEP-2004
DEFINITION Aug 1 Juglans regia cDNA clone WSC0002_IVF_All 5', mRNA sequence.
CV195867
VERSION CV195867.1 GI:52124704
KEYWORDS EST.
SOURCE Juglans regia (English walnut)
ORGANISM Juglans regia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicotyledons;
Rosids; eurosids I; Fagales; Juglandaceae; Juglans.
1 (bases 1 to 766)
Muir, R., Baek, J., Leslie, A., Cook, D. and Dandekar, A.
Analysis of genes expressed in walnut seed coat tissue
Unpublished (2004)
Contact: Abhaya Dandekar, PhD
CAES Genome Facility
UC Davis, Department of Pomology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 752 7784
Fax: 530 752 8502
Email: andandekar@ucdavis.edu
Seq primer: WSCP-TCCGAGATCTGGACGAGC.
FEATURES
source
1..766
Location/Qualifiers
/organism="Juglans regia"
/mol_type="mRNA"
/cultivar="Tulare"
/db_xref="taxon:51240"
/clone="WSC0002_IVF_All"
/sex="Hermaphrodite"
/dev_stage="Mid season fruit collected Aug 1"
/lab_host="XL10-Gold"
/clone_lib="Seed coat from mid-season walnut embryos
collected Aug 1"
/note="Organ: Seed coat; Vector: pTriplex2; Site 1: SfIIA;
Site 2: SfIIB; Walnut nut samples were harvested from
Tulare trees growing in the 'Stuke Block' in the Wolfskill
experimental orchard located in Winters, California (USA)."
```

Samples were harvested on August 1, 2001 between 8 and 10 am. Samples were then dissected the same day. Seed coat (pellicle) tissue was separated from embryos and frozen immediately in liquid nitrogen and stored at -80C. A gram of sample was removed and ground to a fine powder in liquid nitrogen. Total RNA was extracted using the hot borate procedure. Poly A+ RNA was obtained using the Poly(A) Purist kit (Ambion). The cDNA library was constructed using the SMART cDNA library Kit (Clontech). Primary library was en masse evicted and plasmid DNA containing the cDNA library was isolated from the resultant bacterial population. Plasmid DNA was then transformed into ultra competent E coli cells (XL10 Gold; Stratagene). Transformants were plated out on Q-trays (2000 cfu/track), picked using a Qbot and archived in 384 well dishes."

ORIGIN

	Query Match	9.7%	Score 279.6	DB 7	Length 766
	Best Local Similarity	60.3%	Pred. No. 2.8e-58		
	Matches 462	Conservative 0	Mismatches 304	Indels 0	Gaps 0
Qy	1368	GGCGGCACAGTCTCTGTAGATGGCAGCAATGCGCACTACTACTCTCTCTCGGCAATGGAAC	1427		
Db	1	GGGGGACGGGACCTGCTGATGAGGACCGGTGCGACAACTTGTACTCAGCAATATCACAA	60		
Qy	1428	TTGCACCGCATGTTGCGCGCAGCAGAAAGATATTTCTATGAATGTGGGCTTTTTTGGGCAAA	1487		
Db	61	ATGAAGTTAATGCTGCAATCAACTGATGAGATGCCTCTAAATTTGGTTTATCGGGAAAA	120		
Qy	1488	GGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACTAGAACGGGCGGATTTGGTTTT	1547		
Db	121	GGGAACAGTGCNAAAACCTGTATGAATCACTATGAATAATTAGAGCCGGGGCAATGGGACTG	180		
Qy	1548	AAATTGCATGAAGACTGGGGCACAACACCAAGTGCAGTCGATCACTCTTTGAGCGTGGCA	1607		
Db	181	AAGCTGATAGGACTGGGGNACTACTCTCTGCTGCATAGCAATGCTTTGAATGTTGCA	240		
Qy	1608	GATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATAGGCGAGGTTATGTA	1667		
Db	241	GAACAAATATGACATCCAGGTTAATATCCACACGGACACTTTGAATGAATCGGGATTTGTT	300		
Qy	1668	GATGACACCCCTAAATGCAATGAAGGCGGCGCATCCATGCGCTACCACTTAGGGAGCG	1727		
Db	301	GAACATTTCTATGCTGCTATTTAAAGGAAGAACTATTCTATACCTACCACTGAAGGTGCT	360		
Qy	1728	GGTGGAGGACATCACTGATGTTTATCACCATGGCAGCGAGCTCAATATTTACCCCTCC	1787		
Db	361	GGTGGTGGTCACTGCTCCAGATATCATTTAAAGTATGTGCTGTAACATGCTCTGCCATCA	420		
Qy	1788	TCACACACCCCGACTATPTCCCTATACCAATTAATACGGTTGCGAGAACACTTAGACATGCTC	1847		
Db	421	TCTACGAACCCACACGGCCCTTTTACTTTTCAAACACTATAGATGAGCATCTTGACATGCTG	480		
Qy	1848	ATGACATGCCACCACTTAGACAAACGCATCCGGGAGGATTTTACAAATTTTCTCAAAGCCGT	1907		
Db	481	ATGCTGTGCCATCACCTCGATAAGATATTTCCAGAAGACGTGCTTTGCTGTAATCAAG	540		
Qy	1908	ATCCGCCCCGGCTCTATPCGGGCTGAAAGATGTGCTCCATGATATGGGTGTGATCGCGATG	1967		
Db	541	ATAAGGGCTGAACAAATTTGCTGCAGAGGATATTTTGCATGATATGGGGCAATTAGCATC	600		
Qy	1968	ACAAGCTCGAATTCGCAAGCAATGGGCGGTGCGAGCGAAGTGATTCCTCAACTTGGCAG	2027		
Db	601	ATTGCTTCCGATGCACAAGCTATGGGTGCGCATTTGGAGAGGTGATTAGCAAACTTTGGCAA	660		
Qy	2028	ACTGCGGATTAAGATAAAAAAGAAATTTGGTAAAGCTTCCTGAAGATGCCAAAGATAAGCAT	2087		
Db	661	ACTGCCCAAGATGAAGTTTACAAGAGGGTCGATTTGACCTTAGTGACACAGACATGAC	720		
Qy	2088	AATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCG	2133		
Db	721	AATCTTCGTATCAAGCGTTACATTTGCTTAATATACAAATAATCTCTG	766		

RESULT 11
CF826586/c

	Locus
	Definition

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS	JOURNAL
TITLE	COMMENT

FEATURES
SOURCE

```

/organism="Coccidioides posadaei"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIDB361"
/dev_stage="saprobic phase (mycelia)"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadaei saprobic phase cDNA
library, 2 to 4 kb"
/note="Vector: pExpress 1; Site1: Not I; Site2: Eco RV;
Coccidioides posadaei saprobic phase cDNA library, size
fractionated cDNA 2 to 4 kb"

```

ORIGIN

Query Match	9.5%;	Score	274.2;	DB	6;	Length	986;
Best Local Similarity	59.1%;	Pred.	No. 6.6e-57;				
Matches	512;	Conservative	0;	Mismatches	343;	Indels	12;
Gaps	2;						
Qy	1721	GGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATCGGAGCGGAGCTCAATATTC	1780				
Db	986	GGGTGCTGGAGGAGGCCAGCTCCAGATATCATTCGTCGTGAGAAGCCAAACGTCCT	927				
Qy	1781	ACCTCTCTCCACCACCCCACTATTCCTATACCATTAATACGGTTGCGAACACACTTAGA	1840				
Db	926	GCCAGGAGTACGAATCCCACTCGTCCGTATACGGTAATACTTTAGATGAACATCTGGA	867				
Qy	1841	CATGCTCATGACATGCCACCCTAGACAAACCGCATCCGCGAGGATTTACATTTTCTCA	1900				
Db	866	CATGGTAAATGGTCTGCCATCATTTGTGCCAAGATATTTCTGAAAGACGTGGCTTTTGGGA	807				
Qy	1901	AAGCGGTATCCGCCCGGCTCTATCCGGGCTGAAGATGTGCTCCATGATATGSGTGTGAT	1960				
Db	806	AAGCCGGATCCGATCCGAGACAAATGTGTGAGAAAGACGTTCTTCATGACACGGGAGCCAT	747				
Qy	1961	CGCGATGACAAGCTCCGATTCGCAAGCAATGGCGCGTGCAGGCGAAGTGAATTCCTCGAAC	2020				
Db	746	CAGATGTCTATCTCCGACTCTCAAGCTATGGGACGCTGTGGAGAGTGTGTTGTTGGAC	687				
Qy	2021	TTGCGCAGCTCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCCTGAAGAT---	2077				
Db	686	ATGGAACATGACATAAGAAATAAATGGAACGAGGGCGACTCAAGGAAGATGAAGGGAC	627				
Qy	2078	AGATAACGATAAATTTCCGCATTAAGGCTTACATCTCCAAATACACTATCAACCCCGCTTT	2137				

Db	626	GGATTCTGATAAATTTTAGGGTTAAACGGTATATACGCAAGTACACCAATCAACCCCTGCCAT	567
Qy	2138	GACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAAGAGCGCAAGATCGCCGACTTGGT	2197
Db	566	TGCACAGGGGATGGCCCACTATTTGGGAGCGTGAAGTTTGGCAAGACCGCTGATTTGGT	507
Qy	2198	GGTGTGGAATCTCGCCCTTTTTCGGCTTAAACCCAAATCGTGATCAAAAGCGGTATGGT	2257
Db	506	TCTGTGGAATTTTGGCCAACTTTTGGGACTAAACCGAGTATGGTCTTGAAGTCTCGAATGGC	447
Qy	2258	GGTCTTCTCTGAAATGGCGGATTCTAAACGGGCTGTGCCCCACTCCCCAACCGGTTTATTA	2317
Db	446	TGTTCTACGCGCAGATGGTGTGATCCCAATGGCTCTATCCCCACAATCGAGCCTATTATTAT	387
Qy	2318	CCGCGAATGTTTGGGCGATCAGCGCAGGCGAAATTTGACACCGAGCATCACTTTTGGTTTC	2377
Db	386	GAGGCGTATGT-----ACGGTAGCTTTAACCTTAAAGCGCTCAATCATGTTCGTATC	336
Qy	2378	CAAAGTCGCCTATGAAATCGCGTGAAGAAAAAGCTGGCTTAGAGCGCCAAAGTTCTPACC	2437
Db	335	CCAAGCATCCATCAAGCTTGGTATCATCGACAGTTACCACTGAAGAAGCGGATCGAGCC	276
Qy	2438	GGTCAAAAACCTGCGGTAACATCAACAAGAAGACTTCAAGTTCAACGACAAACGGCAAA	2497
Db	275	AGTGAAGAAATTTGTTCGGAATATAAGCAAGAGAGATATGAAATTTTAATGATATTATGCCCAA	216
Qy	2498	AATCACCGTCGATCCGAAAAACCTTTCGAGGTTCTTTGTAGATGGCAAACTCTGCACCTCTAA	2557
Db	215	AATGAGAGTCGATCCGAGAGACTATGTGTGCGGCTGACGGGGAAGATGCACCGCTGA	156
Qy	2558	ACCCACCTCGCAAGTGCTCTTAGCCCCA	2584
Db	155	GCCAGTGTCAAGATTGCTTTTAAACACA	129

RESULT 12	AZ935182/c
LOCUS	AZ935182
DEFINITION	BJ_Ba000302of B. japonicum BAC library Bradyrhizobium japonicum genomic, genomic survey sequence.
ACCESSION	AZ935182
VERSION	AZ935182.1 GI:13777494
KEYWORDS	GSS.
SOURCE	Bradyrhizobium japonicum Bradyrhizobium japonicum
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE	1 (bases 1 to 784)
AUTHORS	Tomkins,J.P., Wood,T.C., Stacey,M.G., Loh,J.T., Judd,A., Goicoechea,J.L., Stacey,G., Sadowsky,M.J. and Wing,R.A.
TITLE	A marker-dense, sequence-ready map of the Bradyrhizobium japonicum genome
JOURNAL	Genome Res. 11 (8), 1434-1440 (2001)
PUBMED	11483585
COMMENT	Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Class: BAC ends High quality sequence stop: 719. Location/Qualifiers 1..784 /organism="Bradyrhizobium japonicum" /mol_type="genomic DNA" /strain="USDAL10" /db_xref="taxon:375" /lab_host="E. coli" /note=Vector: pindigo536; site_1: HindIII"
FEATURES	
source	
ORIGIN	

ORIGIN

Query Match	9.5%	Score 273	DB 9	Length 784
Best Local Similarity	64.0%	Pred. No. 1.2e-56		
Matches 458	Conservative 0	Mismatches 253	Indels 5	Gaps 3
Qy	1468	ATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTTTGTAGAACAAAGTAG	1527	
Db	715	ATCTCGGCATTTTCGGGCAAGGGCAACGCTTCGGCCCGCGCGCTGTCGAGATGATCA	656	
Qy	1528	AAGCGGGCGCGATTGGTTTTAAATTGCATGAAGACTTGGGGCAACAACCAAGTGCATCG	1587	
Db	655	AGNCGGCGCATCGCGCTGAAGCTGCACGAGATTGGGGCAC-ACGCCGGCGCGATCG	597	
Qy	1588	ATCACTGCTTGACGCTGGGAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAG	1647	
Db	596	ACAACCTGCCCTGTCGGTGGCGAGCATTTACGACATCCAGGTCAATGATCCACACCGATACGC	537	
Qy	1648	TCAATGAGGCAGGTTATGTAGATGACACCTTAATGCAATGAACGGGGCGGCCCATCCATG	1707	
Db	536	TGACGGAATCGGGCTTCGTCGAGGATACGATCAAGGGGTTCAAGGGCGCACCATCCACG	477	
Qy	1708	CCTACCAACATTGAGGAGCGGGTGGAGGACATCACTGATGTTATCAACGCGAGGCG	1767	
Db	476	CCTTCCACACCGAGGGCGCGCGGGCTCACGCCCGGACATCATCAAGGTCGCAAGGC	417	
Qy	1768	AGCTCAATATTCACCTCTCTCACCAACCCCACTATTCCCTATACCATTAATACGGTTG	1827	
Db	416	TGAAGAACGTGCTGCGCTCATCGACCAACCGGACGCGCCCTTCACCCGCAACACCATCG	357	
Qy	1828	CAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATT	1887	
Db	356	ACGAGCATCTCGACATGCTGATGGTGTGCCACCACTCGATCCCTCGATCGCGAAGATC	297	
Qy	1888	TACAAATTTTCTAAAGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGCTCCATG	1947	
Db	296	TTGCGTTTCGCGAAGACCGGTATCCGCAAGGAGACCATCGCGGCGGAGCATCTTGCAAG	237	
Qy	1948	ATATGGGTGTATCGCGATGACAAAGCTTCGGATTTCGCAAGCAATGGG-GCGTGCAGGCGAA	2006	
Db	236	ATCTCGGCGGCTCTCGATGATGCTTCGNACTCCAGGCGCATGGGCGCCCTCGGGCGAA	177	
Qy	2007	GTGATTCCTCGAACTTGGCAGACTGGCGGTAAGAAATAAAAAAGAAATTTGTAAGCTTCC-	2065	
Db	176	GTATCATCTCGGACCTGGCAGACGCGCGACAAGATGAAGAGCAGCGGATCGTGCAG	117	
Qy	2066	--TGAAGATGGCAAGATTAACGATAATTTCCGATTTAAGCGCTACATCTCCAATACACT	2123	
Db	116	CAGGACAAGGGCAAGGACACGCAATTTCCGCGTCAAGCGCTACATCGCAAAATACAG	57	
Qy	2124	ATCAACCCCTTTTGACCCACGCGTTCGAGGAGTATATCGGCTCTGTGGGAAGGGG	2179	
Db	56	ATCAACNCCGCGATCGCGACGCGGTGTCGAAGCTGATCGGTTGCGTGGAGAAGGG	1	
RESULT 13				
CO027797/c				
LOCUS	CO027797	990 bp	mRNA	linear EST 10-JUN-2004
DEFINITION	EST806181 Coccidioides posadasii	spherule cDNA library, 0.5 to 5.3		
	Kb Coccidioides posadasii	cDNA clone CIFAM57 3' end, mRNA sequence.		
ACCESSION	CO027797			
VERSION	CO027797.1	GI:48558214		
KEYWORDS	EST.			
SOURCE	Coccidioides posadasii			
ORGANISM	Coccidioides posadasii			
REFERENCE	Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;			
AUTHORS	Onygenales; mitosporic Onygenales; Coccidioides.			
TITLE	1 (bases 1 to 990)			
	Gardner, M.J. and Cole, G.T.			
	Analysis of gene expression in Coccidioides posadasii mycelia and			
	spherules via expressed sequence tags			
JOURNAL	Unpublished (2003)			
COMMENT	Other ESTs: EST806182			
	Contact: Gardner MJ			

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301 838 3519
Fax: 301 838 0208
Email: gardner@tigr.org

FEATURES
source

Location/Qualifiers
1..990
/organism="Coccidioides posadasii"
/mol_type="mRNA"
/strain="C735"
/db_xref="taxon:199306"
/clone="CIFAM57"
/dev_stage="spherules"
/lab_host="E. coli DH10B, T1 phage resistant"
/clone_lib="Coccidioides posadasii spherule cdna library,
0.5 to 5.3 kb"
/notes="Vector: pExpress 1; Site 1: Not 1; Site 2: Eco RV;
Coccidioides posadasii spherule cdna library, 0.5 to 5.3
kb"

ORIGIN

Query Match 9.1%; Score 262.6; DB 7; Length 990;
Best Local Similarity 58.8%; Pred. No. 5.2e-54;
Matches 525; Conservative 0; Mismatches 339; Indels 29; Gaps 3;
QY 1696 GGGCCATCATGCTACCACTTGGAGGCGGTGGAGGACACTCACCTG-ATGTTATC 1754
DB 990 GAGTGATTCATAGTACCACTAGAGGTGCTGGAGAGGCCACGTCAGATATCAT 931
QY 1755 ACCATGGCAGGAGTCAATATTCACCTCTCCACACCCCCCACTATTCCTATACC 1814
DB 930 TCCGTCGTCGAGAGCAAAAGTCTGCCACGACGATCGAATCCCACTGCTCGTATACG 871
QY 1815 ATTAATACGGTTGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGC 1874
DB 870 GTAAATACTTTAGATGAACATCTGGACATGGTAATGCTCTGCCCATCATTTGTCCAAAGAT 811
QY 1875 ATCCGCGAGGATTTACAATTTCTCAAGCGGTATCGGCCGCGCTATCGCGGCTGAA 1934
DB 810 ATCTCTGAAGAGCTGGCTTTTGGGAAGCGGATCCGATCCGAGCAATTTGCTGCAGAA 751
QY 1935 GATGTCTCCATGATATGGGTGTGATCGGATGACAAAGTTCGGATTCGCAAGCAATGGGG 1994
DB 750 GAGTTCTTCATGACACGGGAGCCATCAGCATGCTATCTCGGACTCTCAAGCTATGGGA 691
QY 1995 CQTGAGGCGAAGTATCTGTAATCTGGCAGACTGGGATGAAGTAATAAAGAAATTT 2054
DB 690 CGCTGTGAGAGAAGTTGTTGTCGACATGGAACACTGCACATAAAGAAATAAATGGAACGA 631
QY 2055 GGTAAAGCTTCCTGAAGAT---GGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATC 2111
DB 630 GGGCGACTCAAGGAAGATGAAGGGACGGAATCTGTATAATTTTAGGGTTAAACGGTATATC 571
QY 2112 TCCAAATACATATCAACCCCGCTTTGACCCAGCGGTGAGCAGTATATCGGCTCTGTG 2171
DB 570 ACCAAGTACACCATCAACCTCGCATTTGCACAGGGGATGCCACACTATTGGAGCGGTG 511
QY 2172 GAAGAGGCGAAGATCCCGACTTTGGTGTGTGGAATCTTGCCCTTTTGGCGTAAACCC 2231
DB 510 GAAATTTGGCAAGACCGCTGATTTGTTCTGTGGAAATTTGCCAACTTTGGGACTAAACCG 451
QY 2232 AAAATCGTGATCAAGCGCGTATGGTGTCTTCTCTGAATATGGCGCATTTCTAAGCGGTCT 2291
DB 450 AGTATGGTCTTGAAGTCTGGAATGGCTGTCTCAGCGCAGATGGGTGTATCCCAATGGCTCT 391
QY 2292 GTGCCCACTCCCAACCGTTTATTACCGGAAATGTTGGGACATCACGCGAAGCGGAA 2351
DB 390 ATCCCAACAAATCGAGCTTATATTATGAGCGCTATGTACGCTCTCTCA----- 344
QY 2352 TTTGACACGACATCACTTTTGTTCCTCAAGTGCCTATGAAATGGCGGTGAAGAAAG 2411
DB 343 -----ATCATGTTCTGATCCCAAGCATCCATCAAGCTTGGTATCATCGACGT 296

QY 2412 CTGGGCTTAGAGCCGCCAAGTTCTACCGTCAAAAACCTCCGTAACATCACCAAGAGAC 2471
DB 295 TACCACCTGAAGAAGCGGATCGAGCCAGTGAAGATTGTTCGGAATATTAAGCAAGAGAT 236
QY 2472 TTCAAGTTTCAACGACAAACGCGAAATACCCGTCGATCCGAAACCTTTCGAGGTCTTT 2531
DB 235 ATGAATTTATGATATTTATGCCCAATGAGATCGATCCGAGAGCTATGTTGTCGAG 176
QY 2532 GTAGATGCGCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCCCA 2584
DB 175 GCTGACGGGAGAGTGCACCGCTGAGCCAGTGCAGAGTTGCCCTTTAACA 123
CN907928 684 bp mRNA linear EST 07-JUN-2004
LOCUS 030109ABLC001286HT (ABLC) Braeburn cell culture three days after
DEFINITION subculture Malus x domestica cdna clone ABLC001286, mRNA sequence.
CN907928
ACCESSION CN907928.1 GI:48380429
VERSION
KEYWORDS EST.
SOURCE Malus x domestica
ORGANISM Malus x domestica
REFERENCE 1 (bases 1 to 684)
AUTHORS Beuning,L., Bowen,J., Crowhurst,R., Gleave,A., Janssen,B.,
McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
TITLE Hortresearch Apple EST Project
JOURNAL Unpublished (2004)
COMMENT Contact: Gleave,A.
Sequencing Facility
The Horticulture and Food Research Institute of New Zealand Ltd
120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
Tel: 00 64 09 815 4200
Fax: 00 64 09 815 4201
Email: est@hortresearch.co.nz.
Location/Qualifiers
1..684
/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="ABLC001286"
/tissue_type="Cell cultures"
/dev_stage="Exponential phase of growth (three days after
subculture)"
/clone_lib="(ABLC) Braeburn cell culture three days after
subculture"
/note="Vector: pBluescript SK(-); Library sequenced by
Genesis Research & Development"

Query Match 9.1%; Score 261.2; DB 7; Length 684;
Best Local Similarity 61.4%; Pred. No. 1.1e-53;
Matches 419; Conservative 0; Mismatches 263; Indels 0; Gaps 0;
QY 1546 TTAATTTGCATGAAGACTGGGGCAACACCAAGTGCATCGATCCTGTTGAGCGTG 1605
DB 3 TGAAGCTGCATGAGGACTGGGGAAGCACTCCTGCTGCAATCGCAATTTGTTGGCTGTTG 62
QY 1606 CAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTAG 1665
DB 63 GAGATCAATATGACATCCAGGTAAACATCCATACAGACACCTTGAATGAATCTGATTTG 122
QY 1666 TAGATGACACCTTAATGCAATGAACGGCGGCCCATCCATCCCTACACATTTGAGGGAG 1725
DB 123 TAGAGCATCAATTTGCTGCAATTTAAGGAAGAACTATCCATCTTACCACAGTGAAGTG 182
QY 1726 CGGTTGGAGGACATCACTACCTGATGTTTATCCACATGGCAGGCGAGCTCAATATTTACCC 1785
DB 183 CAGGTGGGGGCCATGCTCCAGATATATCAAAAGTCTCGGGTGTGAAAAATGTCTCTGCCAT 242

```
QY 1786 CCTCACCACCCCACTATTCCCTATACCATTAATACGGTTGCAGAACTTAGACATGC 1845
Db 243 CATCAACGAATCCCAACGGCTTTTCCCTCAAAATACCTATAGATGAGCATCTTGATATGC 302
QY 1846 TCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTTACAATTTTCTCAAAGCC 1905
Db 303 TGATGGTTTGCATCACCTTGACAAGGACATTCAGAAGATGTAGCTTTTGTGTAATCA 362
QY 1906 GTATCCGCCCGGCTTATCCGGCTGAAGATGTCTCCATGATATGGGTGTGATCGCGA 1965
Db 363 GAATAAGGGCTGAACAATTTGTCAGAGAATATTTTGCACGATATGGGGCCATTAGCA 422
QY 1966 TGACAAGCTCGATTCGGAAGCAATGGGGCGTGCAGGCAAGTATTCCTCGAACTTGGC 2025
Db 423 TTGTATCTTCTGATTCACAGGCTATGGGTGCGATTTGGAGAGGTGATTAATCAAGAACTTGGC 482
QY 2026 AGACTGCGGATGAAGTAAAGAAATTTGGTAAGCTTCTCCTGAAGATGCAAGATAACG 2085
Db 483 AAACAGCTGACAGAATGAATCACAAGAGGGTCTGATAGAACTAGTGGATCCGCAATG 542
QY 2086 ATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACG 2145
Db 543 ACAACTTCGGATCAACGATACATCGCAAGTACACTATAATCCAGCAATAGTTAATG 602
QY 2146 GCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCCCGACTTTGGTGTGGA 2205
Db 603 GATTTTCTCAGTATGTGGTTCTATTGAGTGGGGAAGTGGCTGATCTGTCTCTATGGA 662
QY 2206 ATCTGCGCTTTTGGCGTAAA 2227
Db 663 AGCCATCGTCTTTGGCGCAAA 684
```

```
RESULT 15
LOCUS CF714848/c
DEFINITION CCFAP343TO C.neoformans strain JEC21 Cryptococcus neoformans var.
ACCESSION CF714848
VERSION CF714848.1 GI:41569007
KEYWORDS EST.
```

```
SOURCE Cryptococcus neoformans var. neoformans (Filobasidiella neoformans
var. neoformans)
Cryptococcus neoformans var. neoformans
Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
Filobasidiella.
```

```
REFERENCE 1 (bases 1 to 921)
Loftus,B.
End sequencing of clones from a Full length enriched, normalized
JEC21 cDNA library
Unpublished (2003)
Other_ESTs: CCFAP343TR
Contact: Brendan Loftus
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: bjoftus@tigr.org
Seq primer: RF.
```

```
FEATURES
source
1..921
/organism="Cryptococcus neoformans var. neoformans"
/mol_type="mRNA"
/strain="JEC21"
/db_xref="taxon:40410"
/clone="CCFAP343"
/clone_lib="C.neoformans strain JEC21"
/note="Vector: pCMVSPORT6; Site_1: NotI_EcoRV; The full
length, normalized library was prepared from a variety of
conditions using RNA provided by Joseph Heitman and
Jennifer Lodge"
```

ORIGIN

```
Query Match 9.1%; Score 261; DB 6; Length 921;
Best Local Similarity 55.4%; Pred. No. 1.3e-53;
Matches 504; Conservative 0; Mismatches 405; Indels 0; Gaps 0;
QY 1694 GCGCGCATCCATCCCTACACATTTGAGGAGCGGGTGGAGGACACTCAGCTGATGTTAT 1753
Db 921 GAGGACTATTCACAGTTACACACTGAAGAGCTGGTGGCGGACATGCTCAGACATCAT 862
QY 1754 CACCATTGGCAGGCGAGCTCAATATTCTACCTCTCCACCACCCCACTATTCCTATAC 1813
Db 861 CGTTGTCTGTGAATATGAATAACGCTTTCCTAGTTCACCAACCCCACTAGGCTTATGC 802
QY 1814 CATTAATACGGTTGCAAGAACACTTTAGACATGCTCATGACATGCCACCACTTAGACAAACG 1873
Db 801 TGTAAACACTCTTGATGAGCATCTTGACATGCTTATGGTCTGTCAACCACTCGATAAGTC 742
QY 1874 CATCCGCGAGGATTTACAATTTTCTCAAGCGGTATCCGCCCGGCTCTATCGCGGCTGA 1933
Db 741 TATCCCGGAGGACATTTGCCCTTTGCCGACTCTCGTATCCGTTCTGAAACCGTTGAGCGGA 682
QY 1934 AGATGTGCTCCATGATATGGGTGTGATCGGATGACAAAGCTCGGATTCGCAAGCAATGGG 1993
Db 681 AGACGTTTACAGGACACAGGCGCGATTTCCATGATCTCATCTGACTGTCAAGCTATGGG 622
QY 1994 GCGTGAGGCGGAAGTGAATTCCTCGAACTTGGCGAGACTGCGGATGAAGATAAAAAGAAAT 2053
Db 621 TCGTATCGGTGAAGTCACTCACTCGTACATGCGCTACGGCTGCGAAGATGAAGCAATTCG 562
QY 2054 TGTGAAGCTTCTGAAGATGGCAAGATAAGATAATTTCCGCATTAAGCGCTACATCTC 2113
Db 561 TGGTCTCTTTGAGGGCGATGAACCCACGAGAGACAACAATAGGGTGAACCGTATGTTGC 502
QY 2114 CAAATACACTATCAACCCCGCTTTTGCCACCGGCTGAGCGAGTATCGGCTCTGTGGA 2173
Db 501 CAAGTACACTATTAACCCCGCAATTTACCACGGTATGTCGCACCTCATCGGTCAAGTTGC 442
QY 2174 AGAGGGCAAGATCGCGACTTGGTGGTGTGAATCTCTGCGCTTTTGGCGGTAAACCCAA 2233
Db 441 CGTGGGCTGTCTGCAGATCTGGTCTTCTGACGGCAGAGTCTCTTGGTCTAGACCGGA 382
QY 2234 AATCGTGTCAAAAGGCGGTATGGTGGTCTTCTCTGAAATGGGCGGATTTAAACGGTCTGT 2293
Db 381 GATGATCTCAAGGGTGGTGTCTATTGTTGGGTGCAATGGGCGGATGCGAACGCTTCGAT 322
QY 2294 GCCACTCCCCAACCGGTTTATTACCGGAAATGTTTGGGCATCACGGCAAGGCGAAAT 2353
Db 321 CCCACTGTGACGCCCGTCAATGGCAGGCCCATGTGGGGCTCTCAGCCTGAGGCGCGTGC 262
QY 2354 TGACACCAGCATCACTTTTGTTCMAAAGTCGCTATGAAATGGCGTGAAGAAAGAGCT 2413
Db 261 ACTCAATTCATTTGTTGGGTGAGCAGGATCTCTTTGACAGGATCTCGTAGAGAT 202
QY 2414 GGGCTTAGAGCGCAAGTTCTTACCGGTCAAAAATCGCGTAACATCAACAAAGAGATT 2473
Db 201 CAACATCAAGAAAGGCGGAGGCTGTCAAGAACTGTGTCGAATTTGAAAGAGGATAT 142
QY 2474 CAAGTTCAACGACAAAACGCAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGT 2533
Db 141 GAAAGTGAATGACAGTATGCCGAAGATGACTGTGTCGATCCGAGACCTACGACGTTCAGC 82
QY 2534 AGATGGCAAACTCTGCACCTCTAAACCCACCTGCAAGTGTCTTAGCCAGCGCTACAC 2593
Db 81 TGAGGGCGCTTTTGGACGCTCCACCGGACAGAAATCCCACTGACCAAGAGATCTT 22
QY 2594 TTTCTTTCTA 2602
Db 21 CGTTTACTA 13
```

Search completed: November 29, 2005, 07:20:55
Job time : 10217 secs

Adj58237 Urease su
Adj58249 Urease xy
Adj58240 Urease xy
Adj58246 Urease xy
Adj58243 Urease xy
Adj58243 Urease xy
Adj37847 H. bizoz
Adj43451 H. pylori
Adj90180 Helicobac
Adj45680 H. felis
Adj75319 Urease ur
Ab55130 Salmonell
Ab55132 Salmonell
Adj12485 DNA encod
Adj37850 H. bizoz
AbA0816 H. felis
AbA1430 H. pylori
AcA34557 Prokaryot
Adu05358 DNA encod
Adu05359 DNA encod

PS Claim 1: SEQ ID NO 1; 76pp; English.

ID ADJ58240 standard; DNA; 2405 BP.
XX AC ADJ58240;
XX DT 06-MAY-2004 (first entry)
XX DE UreaseXY subunit encoding sequence #1.
XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX OS Helicobacter felis.
XX FH Key Location/Qualifiers
XX FT CDS 1..681
XX FT /*tag= a
XX FT /product= "urease protein"
XX FT 692..2398
XX FT /*tag= b
XX FT /product= "urease protein"
XX XX BP1176192-A2.
XX PN 30-JAN-2002.
XX PD
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX XX (ALKU) AKZO NOBEL NV.
XX PA
XX PI Kusters JG, Cattoli G;
XX PI WPI: 2002-124384/17.
XX DR P-PSDB; ADJ58241, ADJ58242.
XX
XX PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX PT the diagnosis of Helicobacter felis infections and in the preparation of
XX PT vaccines.
XX PS Disclosure; SEQ ID NO 4; 76pp; English.
XX
XX CC The present invention relates to a novel Helicobacter felis urease X and
XX CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX CC used in the manufacture of vaccines against Helicobacter felis infections
XX CC and in diagnostic tests to detect antibodies against Helicobacter felis.
XX CC Helicobacter felis is difficult to grow so it is more convenient to use
XX CC the expression products of the genes encoding the urease X and Y subunits
XX CC in the manufacture of vaccines. The present sequence represents an
XX CC ureaseXY subunit of the invention.
XX
XX SQ Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2405;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 206 GTGAACATCAACCCAAAGAGCAAGAAAGTCTTGTATATATATATGCGGCGAAGTGCT 265
DB 1 GTGAACATCAACCCAAAGAGCAAGAAAGTCTTGTATATATATATGCGGCGAAGTGCT 60

QY 266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCCAAGCCATTGCTTACATTAGT 325
DB 61 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCCAAGCCATTGCTTACATTAGT 120

QY 326 GCCCATATATGACGAAGCGCGCTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
DB 121 GCCCATATATGACGAAGCGCGCTGGAAAAAACCCTTGCTGAACTTATGGAAGAA 180

QY 386 TGCATGCACCTTTTGA AAAAAGATGAAGTAATCCCGGGGTGGGTAAATATGTTCCCGAT 445
DB 181 TGTATGCACCTTTTGA AAAAAGATGAGGTGATCCCGGTGTGGGGAATATGGTCCCTGAT 240

QY 446 CTAGGTGTAGAAGCCACCTTCTCTGATGGTACGAACCTTGTAACCTGTGAATTGCCCATC 505

DB 241 TTGGCGGTAGAAGCCACTTTTCCCGATGGCAACAACTCGTAACCGTGAATTTGGCCCAT 300
QY 506 GAAACAGATGAGCACTTCAAAAGCGGGCGAAAGTGAATTTGGTTGGCATAAAGACATCGAG 565
DB 301 GAACTGATGAACACTTTAAAGCCGGTGAAGTGAATTTGGCTGTGATAAAGACATTTAG 360
QY 566 CTCAATCAGCAAGAAAGTAACCGAACTTGAGGTTACTTAATGAAGGGCTTAATCCTTG 625
DB 361 CTCAACGCGGTAAAGAAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATCCTTG 420
QY 626 CATGTGGGTAGCCATTTCCACTTTCTTGAAGCTAACAGGCACTAAAATTCGATCCTGAA 685
DB 421 CATGTGGGTAGCCATTTCCACTTTCTTGAACCAACAAAGGCATGAAATTCGATCGGAA 480
QY 686 AAAGCTATGCAAAACGCTAGATATTCCTCTGCAACACACGCTACGCAATGGGGCAGGA 745
DB 481 AAAGCTATGCAAAACGCTAGATATTCCTCTGCAACACACGCTACGCAATGGGGCAGGA 540
QY 746 CAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAAAAAGTGAATGGCATGAAC 805
DB 541 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAAGTGAATGGCATGAAC 600
QY 806 GGGCTTGTGAATAACATCGCGATGAACGCCATAAACATAAAGCGCTTGCAAGGCGAAA 865
DB 601 GGGCTTGTGAATAATAATTTCGGAACGAAACGCCATAAACACAAAGCACTAGACAAGGCAAAA 660
QY 866 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAATGAAAAACAAGATATCTAAA 925
DB 661 TCTCAGCGATTCATCAAGTAAGGAGACTCCCATGAAATGAAAAATGAAAAACAAGATATCTAAA 720
QY 926 TACCTACGGAACCCCAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 985
DB 721 CACCTACGGAACCCCAAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 780
QY 986 AGTAGAACATGACTATACCACTATGCGAAGAACTTAAATTTGGCGGGTAAACCTAT 1045
DB 781 AGTAGAACATGACTATACCACTATGCGAAGAGCTCAAAATTTGGCGGGTAAACCTAT 840
QY 1046 CCGTAGGGTATGGTCAAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCACTAC 1105
DB 841 CCGTAGGGTATGGTCAAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCACTAC 900
QY 1106 TAAACGCGATGATTATCGACTACACCGGATTTACAAAGCCGACATTTGGGATTAACAAACGG 1165
DB 901 CAAACGCGATGATTATCGACTACACCGGATTTATTAAGCCGACATTTGGTATTAAATATGG 960
QY 1166 CAAAATCCATGGCATTTGGCAAGGCAAGGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCA 1225
DB 961 CAAAATCCATGGTATTTGGCAAGGCGGGGAACAAAGACATGCAAGATGGCGTAAGCCCTCA 1020
QY 1226 TATGGTGTGGGTGGGCACAGAAGCACTAGCAGGGGAAGTATGATTTATACCGCTGG 1285
DB 1021 TATGGTGTGGGTGGGCACAGAAGCACTAGCAGGGGAAGTATGATTTATACCGCTGG 1080
QY 1286 GGGATCGATTTCACACCCACTTCTCTTCTCCCAACAAATTCCTACCGCTCTAGCCAA 1345
DB 1081 GGGATCGATTTCGCAACCCACTTCTCTCTCCCAACAAATTCCTACCGCTCTAGCCAA 1140
QY 1346 TGGCGTTACAAACCATGTTTGGAGGCGGCACAGGTCCTGTAGATGGCAAGATGGCACTAC 1405
DB 1141 TGGTGTACAAACCATGTTTGGAGGTGGCAAGGTCGGTAGATGGCAAGATGGCACTAC 1200
QY 1406 TATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTTGGCGCAGCAGAGATTTCTAT 1465
DB 1201 CATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTTGGCGCAGCAGCTGAAGAGTATTTCTAT 1260
QY 1466 GAAATGTGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGAACAACT 1525
DB 1261 GAAATGTGGCTTTTGGCAAGGCAATAGCTCTAGCAAAAACAACTTGTAGAACAACT 1320
QY 1526 AGAAGCGGCGCGATTTGGTTTTAAATTCATGAAGACTGGGGCAACAAACCAAGTCGAT 1585

Db 1321 AGAAGCGGGCGGATTGGCTTTAAATTGCA TGAAGACTGGGCGACAACACCAAGTCCGAT 1380
QY 1586 CGATCACTGCTTGACGTGGCAGATGAATATCATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1381 CGATCACTGCTTGACGTGGCAGATGAATATCATGTGCAAGTTTGTATCCACACCGATAC 1440
QY 1646 AGTCAATGAGGAGGTTATGTAGATGACACACCTAAATGAATGAACGGGCGGCCATCCA 1705
Db 1441 GGTCAATGAGGAGGTTATGTAGATGACACACCTAAATGCGATGAACGGGCGGCCATCCA 1500
QY 1706 TGCCTACACATTTAGGAGGCGGGTGGAGGACACTCACCTGATGTTATCACCATGCGAG 1765
Db 1501 TGCCTACACATTTAGGAGGCGGGGAGGAGCACTCACCTGATGTTATCACCATGCGAG 1560
QY 1766 CGAGCTCAATATTTCTACCCCTCTCCACACACCCCCACTATTCCCTATACCAATTAATACGGT 1825
Db 1561 CGAGCTCAATATTTCTACCCCTCTCCACACACCCCCACTATTCCCTATACCAATTAATACGGT 1620
QY 1826 TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGGATCCGCGAGGA 1885
Db 1621 TGCAGAACACTTTAGACATGCTCATGACATGCCACCACTAGACAAACGGATCCGCGAGGA 1680
QY 1886 TTTACAATTTTCTCAAAGCCGATATCGGCCCGGCTCTATCGGGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAAGCCGATATCGGCCCGGCTCTATTTGGCGCTGAAGATGTGCTCCA 1740
QY 1946 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGTGCAGCGGA 2005
Db 1741 TGATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGCTGCAGGA 1800
QY 2006 AGTGATTCCTCGAAGCTTGGCAGACTCGCGATGAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTCCTCGAAGCTTGGCAGACTTGCAGACAAAGCAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 1860
QY 2066 TGAAGATGCAAGATGAACGATAATTTCCGATTAAGCGGTACATCTCMAAATACACTAT 2125
Db 1861 TGAAGATGCGAGATGAATGACAACTTCGCGCATCAAAACGCTATATCTCAAATACACCAT 1920
QY 2126 CAACCCCGCTTTGACCCACGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCATGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGAT 1980
QY 2186 CGCCGACTTGTGTGTGAATCCTGCCCTTTTGGCGTAAACCCCAAAATCTGTGATCAA 2245
Db 1981 CGCCGACTTGTGTGTGAATCCTGCCCTTTTGGTGTAAACCCCAAAATCTGTGATCAA 2040
QY 2246 AGCGGTATGTGTGTCTCTGAAATGGCGGATTTAAACGCTCTGTGCGCCACTCCCA 2305
Db 2041 AGCGGTATGTGTGTCTCTGAAATGGCGGATTTCTAACGCTCTGTGCGCCACTCA 2100
QY 2306 ACGGTTTATACCGGAAATGTTGGGCATCAGCGCAAGCCAAATTTGACACCCAGCAT 2365
Db 2101 GCGCGTTTATACCGGAAATGTTGGGCATCAGCGCAAGCCAAATTTGACACCCAGCAT 2160
QY 2366 CACTTTTGTTCCAAAGTCGCCATGAAATGGCGTGAAGAAAGAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTGTTCCAAAGTCGCCATGAAATGGCGTGAAGAAAGAGCTTAGAGCG 2220
QY 2426 CCAAGTTCTACCGGTCAAAACCTGCGTAAACATCACCAGAAAGACTTCAAGTTCACGA 2485
Db 2221 CAAGGTGCTACCGGTCAAAACCTGCGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
QY 2486 CAAAACGGCAAAATCACCCTGATCCGAAACCTTCGAGGCTTGTAGATGCGCAACT 2545
Db 2281 CAAGACGGCGCATATCACTGTGCATCTTAAACCTTCGAGGCTTGTGTAGATGCGCAACT 2340
QY 2546 CTGCACTCTAAACCCCACTCGCAAGTCTCTAGCCCGAGCGCTACACTTTCTTCTAGGC 2605
Db 2341 CTGCACTCTAAACCCCGCTCTGAAAGTGCCTCTAGCCCAACGCTACACTTTCTTCTAGGC 2400
QY 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4
ADJ58246
ID ADJ58246 standard; DNA; 2407 BP.
XX
AC ADJ58246;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit encoding sequence #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
CDS 2..682
FT /*tag= a
FT /product= "urease protein"
FT 693..2399
FT /*tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
XX 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
XX (ALKU) AKZO NOBEL NV.
XX
PI Kueters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR P-PSDB; ADJ58247, ADJ58248.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 10; 76pp; English.
XX
XX The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 2407 BP; 720 A; 574 C; 584 G; 528 T; 0 U; 1 Other;

Query Match 74.9%; Score 2158.6; DB 7; Length 2407;
Best Local Similarity 93.6%; Fred. No. 0;
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 206 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTATATATTATCGCGGCGAAGTGGCT 265
Db 2 GTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGTATATATTATCGCGGCGAAGTGGCT 61
QY 266 AGAAAGCGCAAGAGGAGGCTTAAAGCTCAACCAACCGAAGCCATTTCCTTACATTAGT 325
Db 62 AGAAAGCGCAAGAGGAGGCTTAAAGCTCAACCAACCGAAGCCATTTCCTTACATTAGT 121
QY 326 GCCCATATTATGAGGAAGCGCGCTGGAAAAAAGAAACCGTTGCCAGCTTATGGAAGAG 385
Db 122 GCCCATATTATGAGGAGCGCGCTGGAAAAAAGAAACCGTTGCCAGCTTATGGAAGAG 181
QY 386 TGCATGCACCTTTTGAAGAAAGATGAAGTATGCCGCGGTGGTAAATATGTTCCCGAT 445

Db 182 TGTATGCACCTTTTGAAGAAAGACGAGTGATGCCCGGTGTGGGAAATATGGTCCCTGAT 241
Qy 446 CTAGGTGTAGAAGCCACTTCTCTGTAGTGTACGAAACTTGTAACTGTGAAATTTGGCCCAATC 505
Db 242 TTAGGCGTGAAGACTACTTTTCCCGATGGCACAACTCTGTAACTCGTGAATTTGGCCCAATC 301
Qy 506 GAACCCAGATGAGCACTTCAAAGCGGGGAAAGTGAATTTGGTTGCGATAAAGACATCGAG 565
Db 302 GAACCCGATGAACACTTCAAAGCGGGGAAAGTGAATTTGGCTGTGATAAAGACATTTGAA 361
Qy 566 CTCAATGCAGGCAAGAAAGTAAACCGAACTTGAGGTTACTAATGAAGGCCCTAAATCTCTTG 625
Db 362 CTCAACGCAAGTAAAGAACTTACCGAACTAGAAGTTACCAACGAAGGACCTTAAATCTCTTG 421
Qy 626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACCAAGGCACATAAAATTCGATCGTGAA 685
Db 422 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCCCAAGGCAATTTGAATTCGATCGGGAA 481
Qy 686 AAAGCCCTATGGCAAAACGCTAGATATTTCCCTCTCGCAACACGCTACGCAATTTGGGGCAGGA 745
Db 482 AAAGCCCTATGGCAAAACGCTAGATATTTCCCTCTCGCAACACGCTACGCAATTTGGGGCAGGA 541
Qy 746 CAAACCCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAGTGAATTTGGCATGAAC 805
Db 542 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGCAGTAAAGTGAATTTGGCATGAAC 601
Qy 806 GGCCTGTGTAATACATTCGGGATGAAGCCATTAACATTAAGCGCTTGCAAGCGCAAA 865
Db 602 GGCCTGTGTAATTAATTTGCAGATGAAGCCATTAACCAAGCGCTTAGAAAAAGCAAA 661
Qy 866 TCTCACGGATTTTCAAGTAAAGAGACTCCCATGAAGTAAAGTAAAGTAAAGTAAAGTAAAG 925
Db 662 TCTCACGGATTTTCAATTAAGAGACTCCCATGAAGTAAAGTAAAGTAAAGTAAAGTAAAG 721
Qy 926 TACCTACGGAACCAACGAGCGATAAAGTTCGCTTAGGAGATACCGATCTTTGGGCAGA 985
Db 722 TACCTACGGAACCAACGAGCGATAAAGTTCGCTTAGGAGATACCGATCTTTGGGCAGA 781
Qy 986 AGTAGAATACATATACCATCTATGGCGAAGAACTTAAATTTGGCGGGGTAAACATAT 1045
Db 782 AGTAGAATACATATACCATCTATGGCGAAGAACTTAAATTTGGCGGGGTAAACATAT 841
Qy 1046 CCGTGAGGATCGGTGAGAGCAATAGCCCTGATGAAACACACCTAGATTTAGTCAATCAC 1105
Db 842 CCGTGAGGATCGGTGAGAGCAATAGTCCAGATGAAACACACCTAGATTTAGTCAATCAC 901
Qy 1106 TAAACGCGATGATTTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTTAAAAACGG 1165
Db 902 CAACGCGATGATTTATTTGACTACACCGGGATTTACAAAGCCGACATTTGGCATTTAAAGATGG 961
Qy 1166 CAAAAATCCATGGCAATTTGGCAAGGCAAGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1225
Db 962 CAAAAATCCATGGCAATTTGGCAAGGCAAGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCA 1021
Qy 1226 TATGGTCTGGGTGTGGGCAAGAGCACTAGCAGGGGAAGGTATGATTTATACCGCTGG 1285
Db 1022 TATGGTCTGGGTGTGGGCAAGAGCAATTTAGCAGGGGAAGGTATGATTTATACCGCTGG 1081
Qy 1286 GGGAAATCGATTTACACACCCACTTCTCTTCTCCACAAATTTCCCTACCGCTCTAGCCAA 1345
Db 1082 GGGAAATCGATTTACACACCCACTTCTCTCTCCACAAATTTCCCTACCGCTCTAGCCAA 1141
Qy 1346 TGGCGTTTACAACTGTTTGGAGCGGCACAGGTCTCTGTAGATGGCAAGATTCGACTAC 1405
Db 1142 TGGCGTTTACAACTGTTTGGCGGTGGCAAGGTCCGCTAGATGGCAAGATTCGACTAC 1201
Qy 1406 TATCACTCCGGGCAATTTGGAACCTTTGCAACCGCATGTTGGCGGCAGCAGAGATTTCTAT 1465
Db 1202 CATCACTCCGGGCAATTTGGAACCTTTGCAACCGCATGTTGGCGGCAGCTGAAGATTTCTAT 1261
Qy 1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAAAT 1525

Db 1262 GAATGTGGCTTTTGGGCAAGGCAATAGCTCCAGTAAAAAACAACCTTTGTAGAACAAAT 1321
Qy 1526 AGAAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTCGGGCACAAACCAAGTGGCAT 1585
Db 1322 AGAAGCGGCGCGATTCGGCTTTAAATTTGCATGAAGACTCGGGCACAACTCCAAGTGAAT 1381
Qy 1586 CGATCACTGCTTTAGCGTGGCGAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1645
Db 1382 CGATCACTGCTTTAGCGTAGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC 1441
Qy 1646 AGTCAATCAGCAGCGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGCCATCCA 1705
Db 1442 GGTCAATGAGGCGAGTTATGTAGATGACACCTTGAATGCAATGAACGGCGCGCCATCCA 1501
Qy 1706 TGCCTACCACTTTAGGAGGAGCGGTGGAGGACACTCACCTGATTTTATCACCATGGCAGG 1765
Db 1502 TGCCTACCACTTTAGGAGGAGCGGCGGAGGACACTCACCTGATTTATCACCATGGCAGG 1561
Qy 1766 CGAGCTCAATATTTACCTCTCTCAACCCCGACTATTCCTATACCATTAATACGGT 1825
Db 1562 CGAGCTCAATATTTACCTCTCTCAACCCCGACTATTCCTATACCATTAATACGGT 1621
Qy 1826 TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGCTATCCGCGAGGA 1885
Db 1622 TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTAGATAAACGCTATCCGCGAGGA 1681
Qy 1886 TTTCACATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGGCTGAAGATGTGTCTCCA 1945
Db 1682 TTTCACATTTTCCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTGTCTCCA 1741
Qy 1946 TGATATGGGTGTGATTCGCGATGACAAGCTCGGATTCGAAGCAATGGGCGGTGCAGCGGA 2005
Db 1742 TGATATTTGGCGTGTGATTCGCGATGACAAGCTCGGATTCGAAGCAATGGGCGGCTGCAGGA 1801
Qy 2006 AGTGATTCCTCGAATTCGCGAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCC 2065
Db 1802 AGTGATTCCTCGAATTCGCGAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCC 1861
Qy 2066 TGAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1862 TGAAGATGTGCAAGATAACGAACTTCGCTATCAACCGCTACATCTCCAAATACACTAT 1921
Qy 2126 CAAACCCGCTTTGACCCACGCGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAGAT 2185
Db 1922 TAAACCCGCTCTAACCCATGGGTAAAGCGAGTATATCGGCTCTGTGGAAGAGGCGCAAAAT 1981
Qy 2186 CGCGCACTGTGTGTGGAATTCCTGCTTTTGGCGTAAACCCCAAAATTCGTATCAA 2245
Db 1982 CGCTGATTTGGTGTGTGGAATTCCTGCTTTTGGTGTGAAACCTTAAGATTTGTATCAA 2041
Qy 2246 AGGCGGTATGGTGTCTTCTCTGAATGGGCGATTTCTAACGCTCTGTGCCACTCCCCA 2305
Db 2042 AGGCGGTATGGTGTCTTCTCTGAATGGGCGACTTCAACGCGTCTCGCTTACACTCA 2101
Qy 2306 ACCGTTTATTTACCCGCAAAATGTTTGGGCATCACGCAAGGCGAAATTTTGACACCAAGCAT 2365
Db 2102 GCCGGTTATTTACCCGCAAAATGTTTGGGCATCACGCAAGGCGAAATTTTGACACCAAGCAT 2161
Qy 2366 CACTTTTGTTCCTAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCG 2425
Db 2162 CACTTTTGTTCCTAAAGTCCGCTATGAAAAATGGCGTGAAGAAAAAACTAGGCTTAGAGCG 2221
Qy 2426 CCAAGTTTCTACGGTCAAAAACTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2222 CAAAGTGTACCCGTTGAAAAAACTGCGCAACATCACTAAGAAAGACTTCAAAATTCACAA 2281
Qy 2486 CAAACGCAAAAAATCACCGTCGATCCGAAAACTTCGAGGCTTTTGTAGATGGCAAACT 2545
Db 2282 CAAAGCGGCGATATCACTGTGATCTCTAAAACTTCGAGGCTTTTGTAGATGGCAAACT 2341
Qy 2546 CTGACCTCTTAAACCCCACTCGCAAGTGCCTCTAGCCCCAGCGCTACACTTTCTTAGGC 2605
Db 2342 CTGACCTCTTAAACCCGCTCTGAAGTGCCTCTAGCCCCAGCGCTACACTTTCTTAGGC 2401


```
QY      2606 ACAATG 2611
      |||||
Db      2402 NCAATG 2407

RESULT 5
ADJ58243
ID      ADJ58243 standard; DNA; 2183 BP.
XX
XX
AC      ADJ58243;
XX
XX      06-MAY-2004 (first entry)
XX
XX      UreaseXY subunit encoding sequence #2.
XX
XX      immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX
XX      Helicobacter felis.
XX
XX      Key
XX      CDS
XX      Location/Qualifiers
XX      3..683
XX      /*tag= a
XX      /product= "urease protein"
XX      694..2181
XX      /*tag= b
XX      /product= "urease protein"
XX
XX      EP1176192-A2.
XX
XX      30-JAN-2002.
XX
XX      11-JUL-2001; 2001EP-00202666.
XX
XX      17-JUL-2000; 2000EP-00202565.
XX
XX      (ALKU ) AKZO NOBEL NV.
XX
XX      Kusters JG, Cattoli G;
XX
XX      WPI; 2002-124384/17.
XX      P-PSDB; ADJ58244, ADJ58245.
XX
XX      Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX      the diagnosis of Helicobacter felis infections and in the preparation of
XX      vaccines.
XX
XX      Disclosure; SEQ ID NO 7; 76pp; English.
XX
XX      The present invention relates to a novel Helicobacter felis urease X and
XX      Y subunit polypeptides and immunogenic fragments. The polypeptides are
XX      used in the manufacture of vaccines against Helicobacter felis infections
XX      and in diagnostic tests to detect antibodies against Helicobacter felis.
XX      Helicobacter felis is difficult to grow so it is more convenient to use
XX      the expression products of the genes encoding the urease X and Y subunits
XX      in the manufacture of vaccines. The present sequence represents an
XX      ureaseXY subunit of the invention.
XX
XX      Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;
XX
XX      Query Match      67.6%; Score 1948; DB 7; Length 2183;
XX      Best Local Similarity 93.3%; Pred. No. 0;
XX      Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY      206 GTGAAACTCACCCCAAGAGCAAGAAAGTCTTGTATATATATGCGGGCGAAGTGCT 265
      |||||
Db      3 GTGAAACTCACCCCAAGAGCAAGAAAGTCTTGTATATATGCGGGCGAAGTGCT 62
      |||||

QY      266 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCTTACATTAGT 325
      |||||
Db      63 AGAAAGCGCAAGCAGAGGGCTTAAAGCTCAATCAACCGAAGCCATTGCTTACATTAGT 122
      |||||

QY      326 GCCCATATTATGGACGAAGCGCGCTGGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
      |||||
```

```
Db      123 GCCCATATTATGGACGAAGCGCGCTGGGAAAAAACCCTTGCTGTAACCTTATGGAAGAA 182
      |||||
QY      386 TGCATGCACCTTTTGAAGAAAGATGAAGTAAATCCCGGGTGGGTAAATATGTTCCCGAT 445
      |||||
Db      183 TGTATGCACCTTTTGAAGAAAGATGAGTGATGCCGGTGTGGGAATATGGTCCCTGAT 242
      |||||
QY      446 CTAGGTGTAGAAGCACCTTTCTGTATGGTACGAAACTTGTAACTGTGAATTTGGCCCATC 505
      |||||
Db      243 TTGGCGGTAGAAGCCACTTTCCCGATGGCACCAAACTCGTAAACGCTGAAATTTGGCCCAT 302
      |||||
QY      506 GAACCATGAGCACCTTTCAAGCGGGCGAAGTGAATTTGGTTGGCATTAAGACATCGAG 565
      |||||
Db      303 GAACCTGATGAACACTTTTAAAGCGCGTGAAGTGAATTTGGCTGTGTGATAAAGACATTGAG 362
      |||||
QY      566 CTCAATCAGCGCAAGAGTAACCGAACTTCAGGTACTAATGAAGGCGCTAAATCTCTTG 625
      |||||
Db      363 CTCAACCTGGGTAAAGGAAGTTACCGAGCTTGAAGTTACCAAGGAAGGACCTTAATCTCTTG 422
      |||||
QY      626 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAACCAAGGCACTAAAAATTCGATCGTGA 685
      |||||
Db      423 CATGTGGGTAGCCATTTCCACTTCTTTGAAGCAACCAAGGCACTGAATTCGATCGGAA 482
      |||||
QY      686 AAGCCTATGCGCAAAACGCTTAGATATTCCCTCTCTGCGCAACACGCTACGCAATTTGGGGCAGGA 745
      |||||
Db      483 AAGCCTATGCGCAAAACGCTTAGATATTCCCTCTCTGCGCAACACGCTACGCAATTTGGGGCAGGA 542
      |||||
QY      746 CAACCCGCAAGTGCGAGTTGATTCCTCTCTGCGTGCAGTAAAGAGTGAATTTGGCATGAAC 805
      |||||
Db      543 CAACCCGCAAGTGCGAGTTAATCCCTCTCTGCGGTAGTAAAGAGTGAATTTGGCATGAAC 602
      |||||
QY      806 GGGCTTGTGAATAACATCGCGGATGAACGCGATAAACATAAAGCGCTTTGACAAGGCGAAA 865
      |||||
Db      603 GGGCTTGTGAATAATAATTGCGGACGAAGCCATAAACAACAAGCACTAGACAAGGCAAAA 662
      |||||
QY      866 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAAAGAAATATGTAAA 925
      |||||
Db      663 TCTCAGCGATTTATCAAGTAAGGAGACTCCCATGAAATGAAAAACAAAGAGTATGTAAA 722
      |||||
QY      926 TACCTACGGACCCCAACCAAGGCGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 985
      |||||
Db      723 CACCTACGGACCCCAACCAAGGCGATTAAGTGGCTTAGGAGATACCGATCTTTGGGCAGA 782
      |||||
QY      986 AGTAGAATCATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAAACTAT 1045
      |||||
Db      783 AGTAGAATCATGACTATACCACTATGGCGAAGAGCTCAATTTGGCGGGTAAAAACTAT 842
      |||||
QY      1046 CCGTGAGGGTATGGGTGACAGCAATAGCCCTGATGAAAAACCCCTAGATTTTAGTCAATCAC 1105
      |||||
Db      843 CCGTGAGGGTATGGGTGACAGCAATAGCCCAATAGCCCAATAGCAACCTTAGATTTAGTCAATCAC 902
      |||||
QY      1106 TAACCGCATGATTTATCGACTACACGGGATTTACAAGCGGACATTTGGGATTTAAAAACGG 1165
      |||||
Db      903 CAACCGCATGATTTATCGACTACACGGGATTTATAAAGCCGACATTTGGTATTTAAAAATGG 962
      |||||
QY      1166 CAAATCATCGCATTTGCGCAAGGAGGAAACAAGGACATGCAAGATGGCGTGAAGCCCTCA 1225
      |||||
Db      963 CAAATCATCGCATTTGCGCAAGGCGGGGAAACAAGACATGCAAGATGGCGTGAAGCCCTCA 1022
      |||||
QY      1226 TATGCTCGTGGGTGTGGGCACAGAACGACTAGCAGGGGAAGGTATGATTTATTAACCGCTGG 1285
      |||||
Db      1023 TATGCTCGTGGGTGTGGGCACAGAACGACTAGCAGGGGAAGGTATGATTTATTAACCGCTGG 1082
      |||||
QY      1286 GGGAAATCATGATTCACACCCGACTTCTCTTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1345
      |||||
Db      1083 GGGGATCGATTCGACACCCGACTTCTCTCTCTCCCAACAATTTCCCTACCGCTCTAGCCAA 1142
      |||||
QY      1346 TGGCGTTACACCACTGTTTGGAGGCGGCACAGGCTCTGTAGATGSCACCAATTCGCGACTAC 1405
      |||||
Db      1143 TGGGTGTTACACCACTGTTTGGAGGTGGGCACAGGCTCCGGTAGATGGCAATTCGCGACTAC 1202
      |||||
QY      1406 TATCACTCCGGCGCAAAATGGAACCTTCACCGCATGTTTGGCGGACGAGAGAGTATTTCTAT 1465
      |||||
```

1203	DB	CATCACTCCGGGCAAAATGGAACTTTGCACCGCATGTTGCGCGCAGCTGGAAGATTA'TTCTAT	1262
1466	QY	GAATGTGGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAAAACAACTTGTAGAACAAAGT	1525
1263	DB	GAATGTAGGCTTTTGGGCAAAAGCAATAGTTCTAGCAAAAAACAACTTGTAGAACAAAGT	1322
1526	QY	AGAAGCGGGCGCATTTGGTTTTAAATTGCATGAAGACTGGGGCACAACAACAAGTGGCAT	1585
1323	DB	AGAAGCGGGCGCATTTGGCTTTAAATTGCATGAAGACTGGGGCACAACAACAAGTGGCAT	1382
1586	QY	CGATCACTGCTTGAGCGTGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1645
1383	DB	CGATCACTGCTTGAGCGTGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1442
1646	QY	AGTCAATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCA	1705
1443	DB	GGTCAATGAGCGAGGTTATGTGGATGACACCCCTAAATGCAATGAACGGGCGCGCCATCCA	1502
1706	QY	TGCCTACCACATTTGAGGGAGCGGTTGGAGACACTACCTGATGTTATCACCATGGCAGG	1765
1503	DB	TGCCTACCACATTTGAGGGAGCGGCGGAGACACTACCTGATGTTATCACCATGGCAGG	1562
1766	QY	CGAGCTCAATATTCTACCCCTCTCCACACCCCGCACCTATTTCCCTATACCAATTAATACGGT	1825
1563	DB	CGAGCTCAATATTCTACCCCTCTCCACACCCCGCACCTATTTCCCTATACCAATTAATACGGT	1622
1826	QY	TGCAGAACACTTAGACATGCTCATGATGCCACCACTAGACAAACGCAATCCGCGAGGA	1885
1623	DB	TGCAGAACACTTAGACATGCTCATGACTGCCACCACTAGATAAGCGCATCCGCGAGGA	1682
1886	QY	TTTACAAATTTTCTAAAGCGGTATCGCCCGGCTCTATCGCGCTGAAAGTGTCTCCA	1945
1683	DB	TTTACAAATTTTCTAAAGCGGTATCGCCCGGCTCTATTTCCGCGCTGAGGATGTGTCTCCA	1742
1946	QY	TGATATGGGTGTGATCCGATGACAAGCTCGGATTTGCAAGCAATGGGCGGTGCAGGCGA	2005
1743	DB	TGATATTTGGCGTGATCCGATGACTAGCTCGATTCCGACGCAATGGGGCGCGCTGGGA	1802
2006	QY	AGTGATTCTTCGAACCTTGGCAGACTCGGATAAAGATAAAAAAGAAATTTGGTAAAGCTTCC	2065
1803	DB	AGTGATTCTTCGAAACTTGGCAAACTGCAGATAAGAAATAAAAAAGAAATTTGGTAAAGCTTCC	1862
2066	QY	TGAAGATGGCAAGATACCAATATTTCCGCATTAGCGCTACATCTCCAAATACACTAT	2125
1863	DB	TGAAGATGGTGCAGATAACGACAACTTCCGCATCAAAACGCTATATCTCCAAATACACCAT	1922
2126	QY	CAACCCCGCTTTGCACCAACCGCGTGAAGCAGTATATCGGCTCTGTGGAAAGAGGCGCAAGAT	2185
1923	DB	TAATCCCGCTTTGACCCATGCGCTGAGCAGTATA TCGGCTCTGTGGAAAGAGGCGCAAGAT	1982
2186	QY	CGCCGACTTGGTGGTGGAAATCTGCTTTTTTTTGGCGTAAAAACCCAAAATTCGTGATCAA	2245
1983	DB	CGCCGACTTGGTGGTGGAAATCTGCTTTTTTTTGGCGTGAACCTTAAGATTTGTGATTA	2042
2246	QY	AGCGGTATGGTGGTCTTCTCGAAATGGCGGATTTCTAAACGCTCTGTGCCACTCCCCA	2305
2043	DB	AGTGGCGATGGTGGTCTTCTCGAAATGGCGGATTTCTAAACGCTCTGTGCCACTCCCA	2102
2306	QY	ACCGGTTTATTACCGCGAAATGTTTGGGCATCACGGCAAGCGGAAATTTTGACACCCAGCAT	2365
2103	DB	GC CGGTTTATTACCGCGAAATGTTTGGGCACCAACGGCAAGCGGAAATTTTGACACCCAGCAT	2162
2366	QY	CACTTTGTGTTTCAAAGTCG 2385	
2163	DB	CAC TT TGTG TGTCTCAAGCG 2182	

RESULT 6
ADQ37847
ID ADQ37847 standard; DNA; 8407 BP.
XX
XX ADQ37847;
XX

DT	07-OCT-2004	(first entry)
XX	H. bizozeronii	urease gene cluster, ureABIEFGH.
XX	Urease;	urease gene cluster; urease structural gene;
XX	urease accessory gene;	ureABIEFGH; Helicobacter bizozeronii infection;
KW	antibacterial; gene;	ds.
KW	Helicobacter bizozeronii.	
XX	US2004142343-A1.	
XX	22-JUL-2004.	
PD	12-AUG-2003;	2003US-00639273.
XX	16-AUG-2002;	2002US-0404337P.
PR	(CHAN/) CHANG Y.	
PA	(SIMP/) SIMPSON K W.	
PA	(ZHUJ/) ZHU J.	
XX	Chang Y, Simpson KW, Zhu J;	
PI	WPI; 2004-533502/51.	
XX	GENBANK; AF330621.	
DR	Novel isolated nucleic acid molecule having urease gene cluster, and	
XX	confering on Helicobacter bizozeronii ability to produce urease, useful	
XX	as vaccine for preventing disease in mammals infected by H.bizozeronii.	
PT	Claim 2;	SEQ ID NO 1; 40pp; English.
PT	The invention relates to an isolated nucleic acid molecule conferring on	
XX	Helicobacter bizozeronii an ability to produce urease, where the nucleic	
CC	acid molecule is a urease gene cluster comprising at least one urease	
CC	structural gene and at least one urease accessory gene. The nucleic acid	
CC	molecule is chosen from ureA, ureB, ureG, ureH and ureI. The	
CC	invention also relates to an isolated protein encoded by the nucleic	
CC	acid, a vaccine for preventing onset of disease in mammals infected by H.	
CC	bizozeronii comprising a nucleic acid and a carrier, and an isolated	
CC	antibody or its binding portion raised against the nucleic acid. The	
CC	nucleic acids, proteins and antibodies are useful for vaccinating mammals	
CC	against onset of disease caused by infection of H. bizozeronii, which	
CC	involves administering the sequences. The sequences are useful for	
CC	detecting H. bizozeronii in a sample of tissue or body fluids which	
CC	involves providing a nucleic acid as an antigen, providing an antibody,	
CC	or providing a nucleotide sequence as a probe in a nucleic acid	
CC	hybridisation assay, contacting the sample with the antigen or the probe,	
CC	and detecting any reaction which indicates that H. bizozeronii is	
CC	present in the sample. This sequence represents the H. bizozeronii	
CC	urease gene cluster, ureABIEFGH.	
XX	Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;	
SQ		

Query Match	41.1%;	Score 1184.4;	DB 12;	Length 8407;
Best Local Similarity	68.5%;	Pred. No. 1.5e-235;		
Matches 1681;	Conservative	0;	Mismatches 741;	Indels 32; Gaps 2;
QY	183	AAATTTTAAACAAGGAGTAA	TAGGTGAAACTCACACC	CAAGACGAAGAAAAGTTCTTGT 242
Db	2319	AAATTTGGTAGAAGGAGTTT	TAGGATGAAATTAACCCCT	AAAGAGCTGCGACAAAGCTCATGT 2378
QY	243	TATATTATCGGCGGAAGT	CGCTAGAAAGCGCAAGCAG	AGGGCTTAAAGCTCAACCAAC 302
Db	2379	TGCNATTATCGGCGGAATT	TGGCTTAAAAACGAAAGCA	ATATGGCGTTAAAGCTAAATTTATA 2438
QY	303	CCGAAGCCATTGCTTTA	CATTAGTCCCATATTAT	TGGACGAAGCGCGCGTGGGAAAAAAA 362
Db	2439	CTGAGGCAGTAGCCCTCAT	CAGTGCCCATGTGAT	TGGAAGAAGCCCGTGCAGGTTAAAAAA 2498
QY	363	CCGTTGCCACGCTTAT	TGGAAAGAGTGCAT	CACTTTTTTAAAAAAGATGAAGTAATGCCCG 422

Db	2499	GTGTGGCGGATTTGATGCAAGAGGAGGAGGACACTTCTTAAAGTGATGATGTATGCTATGCCCCG	2558
Qy	423	GGGTGGTAATATGGTTCCTCGATCTAGGTGTAGAAGCCACCTTTCCTGATGGTACGAAAC	482
Db	2559	GTGTAGCCCATATGATCCAGGAAGTGGGATTTGAAGCTAACTTCCCTGATGGGCAAAAC	2618
Qy	483	TTGTAACTGTGAAATTTGGCCCATCGAACCCAGATGAGCAGCTTCAAAGGGGGGGAAGTGAAT	542
Db	2619	TGGTAACCATCATACCCCGGTTGAAGATGGTGGGCATAAATTTGGCTCCGGGTGAAGTGA	2678
Qy	543	TTGGTTGGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAACTTGAGGTTA	602
Db	2679	TTTTGAAAAACGAAGACATCACTTTGAATGCAAGGCAAAACCAAGCCACCTTTAGAAAGTGC	2738
Qy	603	CTAATGAAGGGGCTTAAATCCTTTCGATGTGGGTAGCCATTTCCACCTTCTTTGAAGCTTAA	662
Db	2739	ATAACAAAGGCGATCGCCCCGTGCAAGTGGCTCCCACTTCCACTTCTTTGAAGTGAATA	2798
Qy	663	AGGCACTAAATTCGATTCGTGAAAAAGCCTATGGGCAAAACGCCCTAGATATTCCTCTTGGCA	722
Db	2799	AGCTTTTGGAAATTTGATCGTGAAGAGCCTATGCGCAAGCCCTAGACATTTGCTTCTGGAA	2858
Qy	723	ACACGCTAGCATTTGGGCGAGGACAAACCGGCAAGTGCAGTTGATTCCTCTTGGTGGCA	782
Db	2859	CCGCTGTGGCTTTGAACCCGGTGAGAAAAAAACCGTGGAAATTTGATTCAAATTTGGCGGTA	2918
Qy	783	GTAAAAAAGTGATTTGGCATGAACGGGCTTTGTAATAACATCGCGGATGAACGCCATAAAC	842
Db	2919	ACCAAGCATTTACGGCTTTAACTCTCTTTGTGATCGCAAGCCGATACCTGATGGCAAA	2978
Qy	843	ATPAAAGCGCTTGACAAAGGCGAAATCTCACGGATTT-----	877
Db	2979	AACTTGTCTCTAAACCGCGCAAGAACATGGCTTTTGGTGTGTGAATTTGGGTTGGGATA	3038
Qy	878	----ATCAGTAAGAGAGATCCCATGAAATGNAAAA---ACAGAAATATGTAATACCT	930
Db	3039	AAAAATTAAGNAAGGACATTCGATGAAAAAAATCTCTCGAAAGAAATATGTTCTATGT	3098
Qy	931	ACGGACCCACAAAGGCGATAAGTCGCTTAGGAGATACCGATCTTTGGGCGAAGATAG	990
Db	3099	ATGGACCCACTAGGGCGATTAAGTGAGATTTGGCGATACCGACCTGATCTTAGAAGTCG	3158
Qy	991	AACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAACTATCCGTG	1050
Db	3159	AACATGACTGCACCTTATGTCGAAGAAATTAAGTTTGGTGGCGGTAAAAACCATTCGCG	3218
Qy	1051	AGGGTATGGTCAGAGCAATAGCCCTGTATGAAACACCCCTAGATTTAGTCATCACTAACG	1110
Db	3219	ATGGGATGGCAAAACCAACAGCCCCCAGCAGCCACGAACTCGATCTTTGTGCTCACTAACG	3278
Qy	1111	CGATGATTTACACTACACCGGATTTTACAAAGCCGACATTTGGGATTTAAAGCGCAAA	1170
Db	3279	CCCTGATCGTGATTTACACCGGATTTTAAAGCCGATATTTGGCATTTAAATTTGGCAAA	3338
Qy	1171	TCCATGGCATTTGGCAAGGAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCATATGG	1230
Db	3339	TCCATGGCATTTGGCAAGGAGGCAATTAAGACATGCAAGATGGCGTTTGCAACAATCTTT	3398
Qy	1231	TCGTGGGTGTGGCACAGAGCACTAGCAGGGGAAGGTATGATTTACCGCTGGGGGAA	1290
Db	3399	GGGTGGGCCCTGTACTAGAGGCTTTTGGCGGCTGAAGGGCTGATTTTACAGCTGGTGGGA	3458
Qy	1291	TCGATTTACACACCCACTTCTTCTCCACAAATTTCCCTACCGCTCTAGCCAAATGGCG	1350
Db	3459	TTGACACCCACATCCACTTTATTTCTTCCCAACAAATTTCCCAAGCATTTGGCAGCGGGA	3518
Qy	1351	TTACAAACCATGTTTGGAGGCGGCACAGGTCTCTGTAGTGCAGCAATTTGGCACTACTATCA	1410
Db	3519	TCACAACCATGATTTGGTGGGGGAAGGTCCAGCTGATGGGACTTAACGCACTACCATCA	3578
Qy	1411	CTCCGGGCAAAATGGAACTTGCAACCGCATTTGCGCGCAGCAGAGAGTATTTCTATGAATG	1470
Db	3579	CTCCGGGGCGCTGGAACCTTTAAACCATGTCTCGTGCCTCTGAAGAAATATGCCATGAAT	3638
Qy	1471	TGGCTTTTGGCAAAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAAAGTAGAAG	1530
Db	3639	TGGCTATTTGGGTAAAGGAATGTGTCTTATGAACCTTCCCTGGTGCATCAACTCGAAG	3698
Qy	1531	CGGGCGGATTTGGTTTAAATTTGATGAAGACTGGGGCAACAAACCAAGTCGATCGATC	1590
Db	3699	CTGGAGCCATTTGGCTTTAAATTTCAACGAAGACTGGGGTAGCACACCTGCAACCTATCTACC	3758
Qy	1591	ACTGCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCATACAGTCA	1650
Db	3759	ATTGCTTGATGTGGCTGACAAATACGATGTGCAAGTGGCTATCCACACCATACCTTGA	3818
Qy	1651	ATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGCGCCATTCATGCCT	1710
Db	3819	ATGAAGCGGCTGTGTGAAGACACTTTTGAAGCCATTGCTGGCGCACATATCCACACTT	3878
Qy	1711	ACCACATTTAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGAGC	1770
Db	3879	TCACACTGAAGGTGCTGGTGGCGGCGCGCTCCGATGTCAATTAAGATGTCTGGCGAAT	3938
Qy	1771	TCAATATTTACCTCTCCACACCCCACTATTTCCCTATACCATTAATACGGTTGCAG	1830
Db	3939	TTAACATCTCTCCAGCTTTTACCAACCCCACTTCCTTTCCCGTGNATACAGAACCG	3998
Qy	1831	AACACTTAGACATGCTCATGACATGCCACCACTTAGACAAAACGCATCCGCGAGGATTTAC	1890
Db	3999	AACACATGGACATGTTGATGTTGCCACCACTTTGGATAAAAACATCAAGAAAGATGCC	4058
Qy	1891	AATTTCTCAAAGCGGTATCCGCCCGGCTCTATTCGGCGGTGAAGATGTGCTCMTGATA	1950
Db	4059	AGTTTGTGATTTCTAGGATTTGCGCCCCCAACCATTCGCGCTTGAGGACAAAATCCACGATA	4118
Qy	1951	TGGGTGTGATCGGATGACAGACTCGGATTCGCAAGCAATCGGGGTGACGCGGAAGTGA	2010
Db	4119	TGGGATTTTCTCTATCACCAGCTCTGACTCCCAGCGATGGGCGGTGAGCGAGGTCA	4178
Qy	2011	TTCTCTCGAACTTGGCAGACTCGGATTAAGAAATAAAAAAGAAATTTGGTAAGCTTCTCGAAG	2070
Db	4179	TCACCCGCACTTGGCAACAGCGGACAAAACAAAGAAATTTGGTCGCTTGCCTGAGG	4238
Qy	2071	ATGGCAAGATAACGATAATTTTCGCAATTAAGCGCTACATCTCCAAATPACACTATCAACC	2130
Db	4239	AAAAAGGCGATATGACAACTTCCGCTCAAGCGCTACATTTCCAAATPACACCATCAACC	4298
Qy	2131	CCGCTTGACCCACCGGCTGAGCGATATATCGGCTCTGTGGAAGAGGCGAAGATCGCGG	2190
Db	4299	CCGCTATTGACACCGGCAATTTCTGAATATGTGCGCTCTGTAGAAAGTGGGCAAAATTCGCGG	4358
Qy	2191	ACTTGGTGTGTGAATCTCTGCCCTTTTGGCGTAAAAACCCAAATCGTGATCAAAGGCG	2250
Db	4359	ATTGCTGCTTTGGAGTCTCTGCGTTCTTTGGCATTTAAACCCCAACATGATCATCAAAGGCG	4418
Qy	2251	GTATGGTGTCTTCTCTGAAATGGCGATTTAAACCGGTCTGTGCCACTCCCCAACCGG	2310
Db	4419	GATTCATCGCACTTCTCTCAAATGGCGATGCCAATGCCTCTATCCCACTCCCCAACCGG	4478
Qy	2311	TTTATTAACCGGAAATGTTTGGGCACTACCGCAAGGCGAAATTTTGACACCGACCTACTT	2370
Db	4479	TGATTTAACCGGAAATGTTTGGCCCACTGGTGAAGCCAAATTTTGACCAATATCATCTT	4538
Qy	2371	TTGTTTCCAAAGTCGCTCTATGAAATGGCGTGAAGAAAAAGCTGGGCTTTAGAGCCCAAG	2430
Db	4539	TTGTATCCCAAGTGGCTTTATGAACCGCATTTAAAGAAAGTGGGCTTGCMAAGAGTGG	4598
Qy	2431	TTCTACCGGTCAAAAACTGCCGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGAAAA	2490
Db	4599	TTTTGCCAGTTAAAACTGCGCAACATCACCAAAAAAGACCTCAAAATTAACGATGTTA	4658
Qy	2491	CGGCAAAATCACCGTGCATCCGAAACCTTTTCGAGGTCTTTGTAGATGGCAAACTCTGCA	2550
Db	4659	CCGCACACATCGAAGTCAATCTCTGAAACCTACAAAGTTTAAAGTGGATGGCAAGAGTTA	4718

Db 915 GTAGAACATGACTACACCAATTTATGGCGAAGAGCTTAAATTCGGTGGCGGTAAACCCCTA 974
QY CTTGAGGGGTATGGGTACAGAGCAATAGCCCTGATGAAACACCCCTAGATTAGTTCACCT 1106
Db 975 AGAGAGGCATGAGCCATCTAACACCTCAGCAAGAGAGTTGGATTAAATATCACT 1034
QY 1107 AACGCGATGATTATCGACTACACCGGGATTTACAAAGCCGACATTTGGGATTAATAACCGC 1166
Db 1035 AACGCTTTAACTCGTGGATTACACCGGTATTTATAAAGCGGATATTGGTATTAAGAGATGGC 1094
QY 1167 AAAATCCATGGCATTTGGCAGGAGGAAACAGGACATCGAAGATGGCTAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAAGCGGTAAACAAAGACATCAAGATGGCTTAAACAAAT 1154
QY 1227 ATGGTCTGGGTGTGGGCACAGAAGCACTAGCAGGGGAAAGGTATGATTTATACCGCTGGG 1286
Db 1155 CTTAGCGTAGTCTGCTACTAGAGCTTAGCCGTTAGAGGTTTGTATCGTAACGGCTGT 1214
QY 1287 GGAATCGATTACACACCCACTTCTCTTCTCCACAAATTCCTACCGCTCTAGCCAAT 1346
Db 1215 GGTATTGACACACACATCCACTTCAATTCACCCCAACAAATCCCTACAGCTTTTGAAGC 1274
QY 1347 GCGGTATCAACCATGTTTGGAGCGGCAAGGTCTCTGTAGATGGACGAATGGACTACT 1406
Db 1275 GGTGTAAACAACCATGATTGGTGGTGAACCGGTCTCTGTGAGGCACTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAACCTTGACCGCATGTTGCGCGCAGCAAGAGTATTTCTATG 1466
Db 1335 ATCACTCCAGGCAGAGAAATTTAAATGATGCTCAGAGCGCTGAAGAATATTTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACTGTAGAACAAAGTA 1526
Db 1395 AATTTAGGTTTCTGGCTAAAGGTAAAGCTTCTAACGATCGAGCTTAGCCGATCAAT 1454
QY 1527 GAAGCGGCGCATGTTTAAATTTGATGAAGACTGGGGCACAAACAAAGTGGCATC 1586
Db 1455 GAAGCGGTGCGATTGGCTTTAAATTCACGAAGACTGGGGCACCACTCTCTTGTCAATC 1514
QY 1587 GATCACTGTTGAGCGTGCAGATGATAGTGTCAAGTGTGATTCACACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGCGGCAAAATACGATGTGCAAGTGCCTATCCACACAGACAT 1574
QY 1647 GTCAATGAGCGAGTTATGTAGATCACCCCTAAATGCAATGAACGGCGCGCATCCAT 1706
Db 1575 TTGATGAAGCGGTTGTGTAGAGACATAATGGCTGCTATTGCTGAGCGCACTATGCAC 1634
QY 1707 GCCTACCACTTGAGGAGCGGTGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACTTTCACACTGAAGCGCTGGCGGCGACACGCTCCTGATATTATTAAGTAGCCGT 1694
QY 1767 GAGCTCAATTTCTACCTCTCCACCCACCCCACTATTCCTTATACCATTAAATACGGTT 1826
Db 1695 GAACACAACTTTCTCCGCTTCCACTAACCCCACTCCCTTTCACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGACATGCAACCACTGACAAACGCAATCCGCGAGGAT 1886
Db 1755 GCAGAGCATGACATGCTATTGTTGTGCCCACTTGGATTAAGACATTAAGAAGAT 1814
QY 1887 TTACAATTTTCTCAAGCGGTATTCGCCCGCGGTCTATCGCGCTGAAAGATGTCTCCAT 1946
Db 1815 GTTCAGTTCGCTGATTCAAGGATCCGCGCTCAAAACCATTTGGCGCTGAACACATTTTGCAT 1874
QY 1947 GATATGGGTGATTCGCGATGACAAAGCTCGGATTTCCGAGCAATGGGGCGTGCAGCGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCACAGTCTGACTCTCAAGCGATGGCGGTGTGGGTGAA 1934
QY 2007 GTGATTCCTCGAATCTGGCAGCTCGGATGAAGATAAAGAAATTTGTAGCTTCTCT 2066
Db 1935 GTTATCACCTAGAACTTTGGCAACAGCTGACAAAAACAAAGAAATTTTGGCCCGCTTGAA 1994
QY 2067 GAAGATGGCAAGATAACGATTAATTTCCGATTAAGCGCTTACATCTCCAAATACACTATC 2126

1995 GAAGAAAAAGGCGATTAACGACAACTTCAGGATCAACCGCTACTTGTCTTAATATACCACTT 2054
2127 AACCCGCTTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATC 2186
2055 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTAGGTTTCAGTAGAAGTGGCAAGTG 2114
2187 GCGCACTTGTGTGTGGAATCTCTGCTTTTGGCGTAAACCCAAATCGTGTATCAAA 2246
2115 GCTGACTTGGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAA 2174
2247 GCGGTATGTTGTGCTCTCTGAAATGGGCGATTCTAACCGGTCTGTGCCCACTCCCCAA 2306
2175 GCGGATTCATTGCGTTAAGCCAAATGGGCGATCGGAACGCTTCTATCCCTACCCCAAA 2234
2307 CCCTTTTATTACCGCAAAATGTTTGGGCATCACCGCAAGCGCAAAATTTGACACCAAGCATC 2366
2235 CCCTTTTATTACAGAGAAATGTTTCGCTCATCATGTTAAAGCTAAATACGATGCAACATC 2294
2367 ACTTTTGTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGAGTGGGCTTAGAGCGC 2426
2295 ACTTTTGTCTCAAGCGCTTATGACAAAGGCAATTAAGAAAGAAATTAGGACTTGAAGA 2354
2427 CAAGTTCTACCGGTCAAAAACCTGCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
2355 CAAGTGTTCGCGTAAAAAAATTCAGAAATATCACTAAAAAGACATGCAATTTCAACGAC 2414
2487 AAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
2415 ACTACTGCTCACAATTGAATCAATCTTGAAACTTTACCATGTGTTCTGTGATGGCAAGAA 2474
2547 TGCACTCTAAACCCACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTTCTAGG 2604
2475 GTAACTTCTAAACAGCCAATAAAGTGAGCTTGGCGCAACTCTTTAGCATTTTCTAGG 2532

RESULT 8
AAQ90180
ID AAQ90180 standard; DNA; 2619 BP.
XX
AC AAQ90180;
XX
DT 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX
XX Helicobacter felis urease ureA/ureB operon.
XX
KW Urease; UreA; UreB; vaccine; Helicobacter infection; Urei;
KW heat shock protein; ss.
XX
OS Helicobacter felis.
XX
PH Key Location/Qualifiers
RBS 31..37
FT /*tag= b
FT /note= "ureA Shine-Dalgarno site"
FT CDS 43..756
FT /*tag= a
FT /EC number= "3.5.1.5"
FT /note= "UreA"
FT RBS 756..759
FT /*tag= d
FT /note= "ureB Shine-Dalgarno site"
FT CDS 766..2475
FT /*tag= c
FT /EC number= "3.5.1.5"
FT /note= "UreB"
XX
WO9514093-A1.
XX
XX 26-MAY-1995.
XX
XX 19-NOV-1993; 93WO-EP003259.
XX

PR	19-NOV-1993;	93WO-EP003259.	
XX	(INSP) INST PASTEUR.		
PA	(INRM) INST NAT SANTE & RECH MEDICALE.		
XX			
XX	Labigne A, Suerbaum S, Ferrero R;		
XX			
DR	WPI; 1995-200383/26.		
DR	P-PSDB; AAR74336, AAR74337.		
XX			
PT	Immunogenic composition against Helicobacter infection - also gene		
PT	fragment(s) and protein(s) from Helicobacter urease gene cluster and heat		
PT	shock protein(s).		
XX			
XX	Claim 10; Fig 3; 128pp; English.		
XX			
CC	The sequence encodes urease UreA and UreB proteins, which are components		
CC	of a novel immunogenic composition capable of inducing protective		
CC	antibodies against Helicobacter infection. The composition may include		
CC	the UreA or UreB proteins, a urease- associated heat shock protein		
CC	(AAR74338-39) or the Urei protein (AAR74340). The composition is used to		
CC	prepare a vaccine for humans or animals, especially against H. pylori and		
CC	H. felis. Antibodies against the proteins may be used for treating		
CC	Helicobacter infection, and primers/probes to the DNA sequence may be		
CC	used for detection of Helicobacter infection. (Updated on 25-MAR-2003 to		
CC	correct PN field.)		
XX			
SQ	Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;		
Query Match 36.9%; Score 1062.6; DB 2; Length 2619;			
Best Local Similarity 65.5%; Pred. No. 2.3e-228;			
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2			
Qy	192	ACRAGGAGTAATAGTGTAAGCTCACCCCAAGACGAGAGGCTTAAAGCTCAACCAACCCGGAAGCCA	311
Db	29	ATAAGGAGTTTAGATGAACTAACGCCCTAAGAACTAGACAAAGTTAATGCTCCATTATG	88
Qy	252	CGGCGCAAGTGGCTAGAAAGCGCAAGCAGAGGCGTTAAAGCTCAACCAACCCGGAAGCCA	311
Db	89	CGGCGAGATGGCAGAGAACCGCTTGGCGTGTGTAACACTCAATTACACCGAAGCGG	148
Qy	312	TTGCTTACATTAGTGCCCATATTATGACGAAAGCGCGCGTGGAAAAAACCCTGTGCC	371
Db	149	TCGCGCTCAATTAGCGGGCGTGTGATGGAAAAAGGCGCGTGTATGTAAGCGGTGGCGG	208
Qy	372	AGCTTATGGAAGAGTGATGCACITTTTGAAGAAAGATGAGTAATGCCGGGTGGGTA	431
Db	209	ATTTGATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAAAATGTGATGACGCGGTAGCAA	268
Qy	432	ATATGGTTCCCGATCTAGTGCTAGAAAGCACCTTTCCTGATGGTACGAAACTGTGAACCTG	491
Db	269	GCATGATTCATGAGTGGGATTGAAGCTAATCTCCCGATGGAACCAAGCTTGTAACTA	328
Qy	492	TGAATTGGCCCATCGAACCGAGATGAGCACTTCAAAGCGGGCGAAGTGAATTTGGTTGG	551
Db	329	TCCACACTCGGTAGAGATAATGGCAAATTAGCCCCCGCGAGGCTTCTTAAA--AA	385
Qy	552	ATAAGACATCGAGCTCAATCGAGGCAAGAGTAACCGAAGCTTGAAGTTACTATGAAG	611
Db	386	ATGAGGACATTACTATTAAACCCCGCAAGAAAGCCATTAGCTTGAAGATGAAAAATGAAG	445
Qy	612	GGCCTAAATCCTTCATGTGGTAGCCATTTTCCACTTCTTGAAGCTAACCAAGGCACATAA	671
Db	446	CGGATCGTCTGTGTCAGTGGGATCAGATTTCCACTTCTTGAAGTGAATAGTCTTTGG	505
Qy	672	AATTTCGATCGTGAAGAACCTTATGGCAACGCCCTAGATATTCCCTCTGGCAACACGCTAC	731
Db	506	ACTTCGATCGCGCAAAAGCTTTTGGCAACGCCCTAGACATTCGATCTGGAACAGCGGTGC	565
Qy	732	GCATTGGGGCAGGACAAACCCGCAAGAGTGAGTTGATTCCTCTTGGTGGCAGTAAAAAG	791
Db	566	GCTTTGAACCCGGGAGGAAAAAGTGTGGAACTCAITTGACATCGCGGGATTAAGCGCA	625

Qy	792	TGATTGGCATGAACGGCGCTTGTTGAATAAATCATCGCGGATGAAACGGCCATAAACAATAAAGCGC	851
Db	626	TCTATGGCTTTAAATCTTTTGGTGGATCGCAAGCCGATGCCGATGTGTAAAAAACTACGGCT	695
Qy	852	TTGCAAGGCGAAATCTCACGGATTT-----	877
Db	686	TAAAAACGGCTAAAGAAAAAGGTTTTTGGGTCTGTAAACTGCGGTTGTGAAGCGACTAAAG	745
Qy	878	-----ATCAAGTAAAGGAGACTCCCATGAAATAGAAAAAACAAGAAATATGTAAATACCTACG	933
Db	746	ATAAACAAATAGGAAAAAACCATGAAAAAGATTTCCAAAAAGAAATATGTTCCTATGTATG	805
Qy	934	GACCCACCAAGGCGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC	993
Db	806	GTCCCACTACCGGGGATCGTGTAGACTCGGCGACACTGATTTTGATCTTAGAAGTGGAGC	865
Qy	994	ATGACTATACCACTATGGCGAAGAACTTAAATTTTGGCGCGGTAAAACTATCCGTGAGG	1053
Db	866	ATGATTGCAACCTTATGTGTAAGAGATCAATTTGGGGCGGTAAACTATCCGTGATG	925
Qy	1054	GTATGGGTCAAGCAATAGCCCTGTATGAAACACCCCTAGATTTAGTCATCACTAAACGCGGA	1113
Db	926	GGATGAGTCAAAACCAATAGCCCTAGTCTTATGAAATTTAGATTTGGTGTCTCACTAACGCCC	985
Qy	1114	TGATTAATCGACTACACCGGATTTTAAAGCGACATATGGGATTTAAAAACGGCAAAATCC	1173
Db	986	TCATTTGGACTATACGGGCATTTTAAAGACCGACATTTGGGATTTAAAGACGGCAAGATTG	1045
Qy	1174	ATGGCATTTGCAAGGAGGAAACAAGACATGCAAGATGCGGTAAAGCCCTCATATGGTCG	1233
Db	1046	CAGGCATTTGCAAGGAGGCAATTAAGGACATGCAAGATGCGGTAGATTAATATCTTTTGGC	1105
Qy	1234	TGGGTGTGGGCAACAGAAGCACTAGCAGGGGAAAGGTATGATTTATCCGCTGGGGGAATCG	1293
Db	1106	TAGTCTCTGCTACAGAGGCTTTGGCAGCTTGAGGCTTGATTTGTAACCGCTGTGTGGCATCG	1165
Qy	1294	ATTTCACACCCCACTTCCTTTCTCCAACAATTCCTCAACCGCTCTAGCCAAATGGCGTTA	1353
Db	1166	ATACGCATATTCACCTTTATCTCTCCCAACAAATCCCTACTGTCTTTTGGCAGCGGGTTA	1225
Qy	1354	CAACCATGTTTGAGGGCGCACAGGTCCTGTAGATGGCACGAATGCGACTACTATCACTC	1413
Db	1226	CAACCATGATTTGAGGAGGACAGACCTTCGGGATGGCACGAATGCGACCATCACTC	1285
Qy	1414	CGGGCAAAATGAAACTTTGCACCGCATTTGCGCGCACGAGAAGTATTTCTATGAATGTGG	1473
Db	1286	CCGGAACGGCTAATCTAAAAAGTATGTTGGTGCAGCGCGAAGNATACGCCATGAATCTAG	1345
Qy	1474	GCTTTTGGGCAAAAGGCAATAGCTCTAGCAAAAAACAATTGTAGAACAAAGTAGAGCGGG	1533
Db	1346	GC'TTTTGGCTAAGGCGAATGTGTCTTACGAAACCTCTTTACGCGATCAGATTTGAAGCAG	1405
Qy	1534	GCGCGATTGGTTTTTAATTTGCATGAGACTGGGGCACACACCAAGTGGGATCGATCACT	1593
Db	1406	GCGCGATTGGTTTTTAATTTCCAGAGACTGGGGGAGCACACTGCACTATTTCACCACT	1465
Qy	1594	GCTTTGAGCTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG	1653
Db	1466	GCCTCAATGTCCCGATGAATACGATGTGCAAGTGTCTATCCACACCGATACCCTTAACG	1525
Qy	1654	AGGCAGGTTATGTAGATGACACCCCTAAAATGCAATGACGGGCGGCCATTCATGCTTACC	1713
Db	1526	AGCGGGCTGTGTAGAAGACACCCCTAGAGCGGATTTCCGGGCGCACCATCCATACCTTCC	1585
Qy	1714	ACATTTGAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGGCAGGCGAGCTCA	1773
Db	1586	ACACTGAAGGGGCTGGGGTGGACACGCTCCGATGTTATCAAAATTTGGCAGGGGAATTTA	1645
Qy	1774	ATATTTTACCCTCTCCACCCACCCCACTATTCCCTATACCATTAATACGGTTTGCAGAAC	1833
Db	1646	ACATTTACCCGCTCTACTAACCCGACCAATCTCTTCCACCAAAACACTGAAGCCGAGC	1705
Qy	1834	ACTTAGACATGCTCATGACATGCCCCACCTAGACAAACGCATCCCGGAGGATTTTACAAT	1893

Db	1706	ACATGACATGTTAAATGGTGTGCCACCACTTGGATAAAAGTATCAAGGAAGATGTGCAGT	1765
Qy	1894	TTTCTCAAAGCCGATATCCGCCCCGGCTCTATCGCGCTGAAGATGTGCTCCATGATATGG	1953
Db	1766	TTGCGGATTCGAGATTTCGCCCCAAACTATCGCGCTGAAGACCAACTCCATGACATGG	1825
Qy	1954	GTGTGATTCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGCTGCAGGCGAAAGTGATTC	2013
Db	1826	GGATCTTTTCTATCACCAGCTCCGACTCTCAGGCTATGGACGCTAGGCGAGGTGATCA	1885
Qy	2014	CTCGAACTTGGCGAGACTGCGGATTAAGATTAATAAAGAAATTTGGTAAAGTTCCTCGAAGATG	2073
Db	1886	CACGCACTTGGCAGACAGCAGACAAAAACAAAAAAGAGTTTGGGCGCTTGAAGAAGAAA	1945
Qy	2074	GCAAGATAACGATAATTTCCGCGATTAAGCGCTACATCTCCAAATACACTATCAACCCCG	2133
Db	1946	AAGCGGATAACGACAACTTCGCGATCAACGCTACATCTCTAAATACACCATCAACCCCG	2005
Qy	2134	CTTTGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCCGACT	2193
Db	2006	GGATCGCGCATGGATTTCTGACTATGTGGCTCTGTGGAAGTGGGCAATAACGCGACC	2065
Qy	2194	TGTTGGTGTGGAAATCTGCGCTTTTGGCGGTAAAAACCAAAATCTGTGATCAAAAGGCGGTA	2253
Db	2066	TCGTGCTTTGGAGTCCGGCTTTCTTTGGCATTAAGCCCAATATGATTTATTAAAGGGGGAT	2125
Qy	2254	TGTTGGTCTTCTGGAATGGGCGATTTCTAAACGGCTGTGCGCCACTCCGCCCAACCGGTTT	2313
Db	2126	TTATTGGGCTCTCTCAAAATGGGCGATGCCAATGCGTCTATTCCACCCCTCAGCCCGTCT	2185
Qy	2314	ATTACCCGCAAAATGTTTGGGCGATCACGGCAAGCGAAATTTGACACAGCATCACTTTTTCG	2373
Db	2186	ATTACCGTGAATGTTTGGACACCATGGGAAAAACAATTCGACACCAATATCACTTTCG	2245
Qy	2374	TTTCCAAAGTCGCTATGAATAATGGCGTGAAGAAAAAGCTGGGCTTAGAGCGGCAAGTTC	2433
Db	2246	TGTTCCCAAGCGGCTTACAAGCGAGGATCAAGGAAGAACTAGGCGTAGATCGCGCGCAC	2305
Qy	2434	TACCGGTCAAAAATGCGCGTAAATCATCAACGAAGAGACTTCAAGTTCACACGACAAACGG	2493
Db	2306	CGCCAGTGAAAAAATGTGCGCAATATCTATAAAGAGGACCTCAAAATTCACGATGTGACCG	2365
Qy	2494	CAAAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCT	2553
Db	2366	CACATATTGATGTCAACCTTGAAACCTATTAAAGTGAAGTGGATGGCAAGAGGTAACTT	2425
Qy	2554	CTAAACCCACCTGCAAGTGCCTTAGCCCAACGCGCTACATTTCTTCTAGG	2604
Db	2426	CTAAAGCAGCAGATGAATTGAGCCTAGCGCAACTTTATTAATTTGTTCTAGG	2476
RESULT 9			
AAT45680			
ID	AAT45680 standard; DNA; 2619 BP.		
XX	AAT45680;		
AC			
XX			
DT	16-OCT-2003 (revised)		
DT	07-MAR-1997 (first entry)		
XX			
DE	H. felis structural urease gene.		
XX			
KW	Urease; UreA; UreB; heat shock protein A; HspA; vaccine; immunogen;		
KW	antigen; antibody; chronic gastritis; ulcer; pIL205; ss.		
XX			
OS	Helicobacter felis; strain ATCC 49179.		
XX			
FH	Key		
RBS	Location/Qualifiers		
FT	31..36		
FT	/*tag= a		
FT	43..753		
FT	/*tag= b		

FT		/standard_name= "URE A"	
FT		/product= "urease subunit A"	
FT	RBS	756..759	
FT		/*tag= C	
FT	CDS	766..2475	
FT		/*tag= d	
FT		/standard_name= "URE B"	
FT		/product= "urease B subunit"	
XX	W09634624-A1.		
XX	PN		
XX	PD		
XX	PF	07-NOV-1996.	
XX	PP	02-MAY-1996;	96WO-EP001834.
XX	PR	19-MAY-1995;	95US-00432697.
XX	PR	19-MAY-1995;	95US-00447177.
XX			
XX	PA	(INSP) INST PASTEUR.	
XX	PA	(INRM) INST NAT SANTE & RECH MEDICALE.	
XX	PI		
XX	PI	Labigne A, Suerbaum S, Ferrero RL, Thiberge J;	
XX	XX		
DR	DR	WPI; 1996-505900/50.	
DR	DR	P-PSDB; AAW06729, AAW06730.	
XX			
XX	PT	New immunogenic compsn. contg. UreB and HspA antigens of Helicobacter -	
XX	PT	for treatment and prevention of esp. H pylori infection, also new	
XX	PT	antibodies specific for these antigens.	
XX	XX		
XX	XX	Example 1; Page 96-100; 184pp; English.	
XX	XX		
CC	CC	A portion (AAAT45680) of plasmid pILL205 (CNMC I-1355) comprises the	
CC	CC	Helicobacter felis structural urease gene encoding urease subunits A	
CC	CC	(AAW06729) and B (AAW06730). pILL205 was obt'd. by examining the	
CC	CC	expression of urease activity by H. felis cosmid clones and	
CC	CC	identification of the H. felis genes required for urease expression when	
CC	CC	cloned in E. coli. The urease gene can be used to design probes for the	
CC	CC	detection of Helicobacter infection, or utilised in the prodn. of	
CC	CC	recombinant urease subunits for use in novel immunogenic compsns.	
CC	CC	(Updated on 16-OCT-2003 to standardise OS field)	
XX	XX		
SQ	SQ	Sequence 2619 BP; 772 A; 569 C; 665 G; 613 T; 0 U; 0 Other;	
	Query Match	36.9%; Score 1062.6; DB 2; Length 2619;	
	Best Local Similarity	65.5%; Pred. No. 2.3e-228;	
	Matches 1606; Conservative	0; Mismatches 804; Indels 41; Gaps 2;	
Qy	192	ACAAGGAGTAATAGTGMAAACTCACACCCAAAGAGCAAGAAAAGTTCTTGTTATATTATG 251	
Db	29	ATAAGGAGTTAGGATGAACCTAACGCTTAAGAACTAGACAGTTAATGCTCCATTATG 88	
Qy	252	CGGCGCAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCGAAGCCA 311	
Db	89	CGGCGCATTTGGCAGAAAGACGCTTGGCGGTGGTGTGAACCTCAATTCACCAAGCGG 148	
Qy	312	TTGCTTACATTAGTCCCATATTATGACCAAGCGCGCTGGAAAAAAACCGTTGCC 371	
Db	149	TCGCGCTCATTAGCGGGCGTGTGATGAAAGCGCGCTGATGTTAATAAAGCGTGGCGG 208	
Qy	372	AGCTTATGGAAGAGTGCATGCATCTTTTGTAAAAAAGATGAAGTAATGCCCGGGTGGGTA 431	
Db	209	ATTGTATGCAAGAAGCGCAGACTTGGCTTAAAAAAGAAAATGTGATGACGCGGTAGCA 268	
Qy	432	ATATGGTTCGGATCTATAGGTGTGAAGCCACCTTTTCTGTATGGTACGAAACTGTGTA 491	
Db	269	GCATGATTCATGAAGTGGGATTTGAAGCTAACTTCCCGGATGGAAACCAAGCTTGTAACTA 328	
Qy	492	TGATTTGGCCATCGAACCAAGATGAGCAGCTTCAAGCGGGCGAAGTGAAATTTGGTTGCG 551	
Db	329	TCCACATCTCGGTAGAGGATTAATGGCAAAATTAGCCCCCGCGAGGCTCTTCTTAA--AA 385	
Oy	552	ATAAGACATCGAGCTCAATTCGACGGCAAGAAAGTAACCGAACTTGAGGTTACTAATGA 611	

Db 386 ATGAGACATTTACTATTAACGCGCGCAAGAACCCATTAGCTTGAAGAGTGAATAAAG 445
Qy 612 GGCCTAAATCTCTTGCACTGGGTAGCATTTCACATTTCTTTGAAGCTTAAACAAGCATAA 671
Db 446 GCGATCGTCTGTGAGGTGGGATCAATTTCCATTTCTTCGAAGTGAATAAGCTTTGG 505
Qy 672 AATTCGATCGTGAAGAGCCTTAGGCAACCGCTAGATATTCCTCTTGCCAAACACGCTAC 731
Db 506 ACTTCGATCGCGCAAAAGCTTTTGGCAACCGCTAGACATTGCATCTGGAACAGCGGTGC 565
Qy 732 GCATTTGGGAGAGCAAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGGCAGTAAAAAAG 791
Db 566 GCTTTGAACCCGCGGAGGAAAAAGTGTGGAATCTATTGACATCGCGGGAATAAAGCGCA 625
Qy 792 TGAATTCGCGATGACCGGCTTGTGAATAACATCGCGGATGAACGCCATATAACATATAAGCGC 851
Db 626 TCTATGGCTTTAAATCTTTTGGTGGATCGCAAGCCGATGCCGATGGTAAAAAACCTCGGCT 685
Qy 852 TTGACAAAGCGAAATCTCAGGATTT----- 877
Db 686 TAAACCGCGCTAAAGAAAAAGTTTGGGTCTGTAACTCGCGTTGTGAAGCGACTAAAG 745
Qy 878 ----ATCAAGTAAGGAGACTCCCATGAAATAAGAAAAAACAAGAAATATGTAAATACCTACG 933
Db 746 ATAAACAATAAGGAAAAACCATGAAAAAGATTTACGAAAAAGAAATATGTTTCTATGTATG 805
Qy 934 GACCCACCAAGCGATAAGTTCGCTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC 993
Db 806 GTCCCACTACCGGGATCGGTGTAGACTCGCGACACTGATTTGATCTTTAGNAGTGGAGC 865
Qy 994 ATGACTATACCACTATGCGGAAGAACTTAAATTTGCGCGGCTTAAACTATCCGTGAGG 1053
Db 866 ATGATTGCACCATTTATGTGAAGAGATCAAAATTTGGGGCGGTAAAACTATCCGTGATG 925
Qy 1054 GTATGGGTGAGAGCAATAGCCCTGTATGAAACAACCCCTAGATTAGTCATCACTAACGCGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTTAGATTGGTGTCTCACTAAACGCC 985
Qy 1114 TGAATATCGACTACACGGGATTTACAAAGCGACATTTGGGATTAATAAGCGCAAAATCC 1173
Db 986 TCAATTTGGACTATACGGGCAATTTACAAAGCCGACATTTGGGATTAATAAGACGGAAGTTG 1045
Qy 1174 ATGGCAATTCGACGAGCAAGAAACAAGCATGCAAGATGCGTAAGCCCTCATATGTCG 1233
Db 1046 CAGGCATTTGGCAAGGCAAGCAATAAGACATGCAAGATGCGTAGATAATAATCTTTTGGC 1105
Qy 1234 TGGGTGTGGCCACAGAAGCACTAGCAGGGGAAGGTATGATTAATACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCTCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTGTAAACGCTGTGGCATCG 1165
Qy 1294 ATTACACACCCACTTCTTTCTCCACAACAATTCCTACCGCTTAGCCAAATGGGGTTA 1353
Db 1166 ATACGCATATTCACTTTATCTCTCCCAACAATACTCCTACTGTCTTTGCGAGCGGGTTA 1225
Qy 1354 CAACCATGTTTGAGCGGCGACAGGTCTGTGTAGATGGCAGCAATGCGACTACTATCACTC 1413
Db 1226 CAACCATGATTTGAGGAGGACAGGACCTTGGGATGGCAGATGGCAGCCACCATCACTC 1285
Qy 1414 CCGGCAATTTGGAATTTGCAACCGCATCTTGGCGCAGCAGAAAGATTTCTATGAATGTGG 1473
Db 1286 CCGAGCGCGTAACTCTAAAAAGTATGTTGGTGACCGGAGAAATACGCCATGATCTAG 1345
Qy 1474 GCTTTTGGGCAAGGCAATAGCTCTTAGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAACCCCTCTTTACGCGATCAGATTTGAAGCAG 1405
Qy 1534 GCGCGATTTGGTTTAAATTTGATGAAGACTGGGGCAACAACCAAGTGCATCGACT 1593
Db 1406 GGGCGATTTGGTTTAAATTTCAAGAGACTGGGGGAAGCACCCTGCGAGCTATTTCACCACT 1465
Qy 1594 GCTTGAGCGTGGCAGATGAATGATGTCAGTTTGTATCCACCGGATACAGTCAATG 1653

Db 1466 GCCTCAATGTGCGCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCTTTAACG 1525
Qy 1654 AGGCAAGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGGCCCATTCATGCTACC 1713
Db 1526 AGCGGGCTGTGTAGAAAGACACCCCTAGAGGCGATTGCCGGCGCACCATTCATACCTTCC 1585
Qy 1714 ACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTTATCACCATGCGAGCGGAGCTCA 1773
Db 1586 ACATCTGAAGGGCTGGGGGTGGACACGCTCCAGATGTTTATCAAAATGCGAGGGGAATTTA 1645
Qy 1774 ATATTTACCTCTCCACACCCCACTATTCCTCTATACCATTAATACGTTTGCAGAAC 1833
Db 1646 ACATTTACCGGCTCTACTAACCCTGACATTCCTTTCCACAAAACACTGAAGCCGAGC 1705
Qy 1834 ACTTAGACATCTCATGACATGCCACCACTAGACAAACGCGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGGACATGTTAATGGTGTGCCACCACTTTGGATAAAAGTATCAAGGAAGATGTGCAGT 1765
Qy 1894 TTTCTCAAGCCGATTCGCGCCCGGCTCTATCGGGGCTGAAGATGTGCTCCATGATATGG 1953
Db 1766 TTGCCGATTCGAGGATTCGCCCCCAACTATCGGGCTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCGCGATCACAAGCTCGGATTCGCAAGCAATGCGGCGTGCAGCGAAGTGATTC 2013
Db 1826 GGATCTTTTCTATCACCAGCTCCGACTCTCAGGCTATGGGACGCTAGCGGAGGTGATCA 1885
Qy 2014 CTCGAACCTTGCGACACTGCGGATAAGATAAAAAAGAAATTTGGTAAGCTTCCTGAAGATG 2073
Db 1886 CACGCACTTTGGCAGACGACAGACAAAACAANAAGATTTTGGCGCTTTGAAGAGGAAA 1945
Qy 2074 GCAAGATAACGATAATTTTCGCAATTAAGCCTACATCTCCAAATPACACTATCAACCCCG 2133
Db 1946 AAGSCGATAACGACAACTTCGCGCATCAACGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGAAGAGGCGAAGATCGCGACT 2193
Db 2006 GGATCGGCGATGGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGTGTGAATCTCGCCCTTTTGGCGTAAACCCAAATCTGTATCAAAAGGCGGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGGCTTTCTTTGGCATTAAGCCCAATATGATTTAAGGGCGGAT 2125
Qy 2254 TGGTGTCTTCTCTGAAATGGCGGATTTAAAGCGGTCTGTGCCCACTCCCCAAACCGGTTT 2313
Db 2126 TTATTTGGCTCTCTCAATGGCGGATGCCAATGGCTCTATTTCCACCCCTCAGCCCGTCT 2185
Qy 2314 ATTAACCGCAAAATGTTTGGGCAATCACGGCAAGGGAATTTGACACCGAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGACACCATGGGAAAAACAAATTCGACACCAATATCACTTTCCG 2245
Qy 2374 TTTCCAAAGTCCCTATGAAAATGCGGTGAAGAAAAGCTGGGCTTAGAGCGCAAGTTTC 2433
Db 2246 TGTCCCAAGCGGCTTACAAGCGAGGATCAAAAGAAAGAACTAGGGCTAGATCGCGGGCAC 2305
Qy 2434 TACCGGTCAAAAACCTGCGGTAAACATCAACAAAGAACTTCAAGTTCAACGCAAAAACCG 2493
Db 2306 CGCAGTGAAAAACCTGTGCAATATCACTAAAAGGAACTCAAAATCAACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTCGATCCGAAAAACCTTTGAGGTCTTTGTAGATGGCAAACTCTGCACCT 2553
Db 2366 CACATATTGATGTCAACCTCGAAACCTATAAGGTGAAGTGGATGGCAAGAGGTAACCT 2425
Qy 2554 CTAAACCCACTCTCGAAGTGCCTTAGCCGAGCGCTACACTTTCTCTTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTTGAGCCTAGCGCAACTTTATAATTTTGTCTAGG 2476

RESULT 10

AAQ75319

ID AAQ75319 standard; DNA; 2619 BP.

XX

AC AAQ75319;

XX

Db 2114 TGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATCATCAATCAAGCGGATTCATTGGC 2173
Qy 2262 TTCTCTGAAATGGCGATTTCTTAACGGCTGTGTCCTCCACTCCCAACCGGTTTATTACCGC 2321
Db 2174 TTAAGCCAATGGCGATGGGAACGCTTCTATCCCTACCCCAACCGGTTTATTACAGA 2233
Qy 2322 GAAATGTTGGGCATCACGGCAAGCGAAATTTGACACGAGCATCTTTTGTTCGAA 2381
Db 2234 GAAATGTTGGCTCATCATGTAAGCTAAATACGATGCAACATCATCTTTGTGTCTCA 2293
Qy 2382 GTCGCTATGAAATGGCTGAAGAAAGCTGGCTTAGAGCGCCAAAGTTCTACGGTC 2441
Db 2294 GCGGCTTATGACAAAGGCATTAAGAAAGATTAAGGACTTGAAGAGCAAGTGTTCGCGTA 2353
Qy 2442 AAAAATCGCGTAAACATCACCAAGAAAGACTTCAAGTTCAACACAAAACGCAAAATC 2501
Db 2354 AAAAATGCGAATAATCACTAAAAAGACATCAATTCACGACTACCGTCAAT 2413
Qy 2502 ACCGTCGATCCGAAACCTTCGAGGCTTTGTAGATGGCAAACTCTGCACCTCTAAACC 2561
Db 2414 GAAGTCAATCTGAACTTACCATGTTCCATGTTCTGATGCGCAAGAAAGTAACTTCTAAACCA 2473
Qy 2562 ACTCGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 2474 GCCAATAAAGTGAAGTGGCGCAACTCTTAGCATTTTCTAGG 2516

RESULT 12
AEB55112
ID AEB55112 standard; DNA; 4824 BP.
XX AEB55112;
AC AEB55112;
XX 22-SEP-2005 (first entry)
XX Salmonella vector pHUR3 expressing helicobacter antigens, SEQ ID 1.
XX Circular; pHUR3; ds; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; immunostimulant; antiulcer;
KW cytotatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.
XX Helicobacter pylori.
OS Salmonella typhimurium.
OS Escherichia coli.
OS Synthetic.
XX WO200132014-A2.
PN 10-MAY-2001.
XX 01-NOV-2000; 2000WO-US030191.
PF 01-NOV-1999; 99US-00431705.
PR (ORAV-) ORAVAX INC.
PA Kleanthous H, Londono-Arcila P, Freeman D;
XX WPI: 2001-343379/36.
DR P-PDB; AEB55113, AEB55114, AEB55115, AEB55116, AEB55117, AEB55118,
DR AEB55119, AEB55120, AEB55121, AEB55122, AEB55123, AEB55124, AEB55125,
DR AEB55126, AEB55127, AEB55128, AEB55129, AEB55131, AEB55132, AEB55133,
DR AEB55134, AEB55135, AEB55136, AEB55137, AEB55138, AEB55139, AEB55140,
DR AEB55141.
XX Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX Disclosure; SEQ ID NO 1; 63pp; English.
XX The invention relates to inducing an immune response against Helicobacter

CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated salmonella vector containing a nucleic acid molecule encoding
CC a Helicobacter antigen, and parenterally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an htrA or
CC nlrB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli AmpR gene.
CC The present sequence represents the pHUR3 plasmid. NOTE: pHUR3 is
CC represented by both AEB55112 and AEB55130, the peptides expressed by
CC pHUR3 are shown in the sequence listing to be split between AEB55112 and
CC AEB55130, yet Figure 4 shows all the peptides being expressed by the one
CC sequence, therefore all encoded peptides are cross-referenced to both
CC sequences.
XX
SQ Sequence 4824 BP; 1389 A; 1090 C; 1146 G; 1199 T; 0 U; 0 Other;
Query Match 36.7%; Score 1058.6; DB 5; Length 4824;
Best Local Similarity 65.7%; Pred. No. 2.2e-227;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;
Qy 197 GAGTATAGGTGAACTCACACCCAAAGAGCAAGAAAGTTCTTTATATATATCGGGC 256
Db 77 GAGATCTCCATGAACTCACCCCAAAAGAGTTAGATAGTTGATGCTCCACTACCTGGA 136
Qy 257 GAATGCTAGAAAGCCCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCCATTGCT 316
Db 137 GAATGCTAGAAAGCCCAAGCAGAGGGCTTAAAGCTCAACCAACCGAGCAGTAGCT 196
Qy 317 TACATTAGTCCCATATTTATGACGAGCGCGCTGGAAGAAACCGTTGCCAGCTT 376
Db 197 TTGNTAGTCCCATATTTATGGAAGAGAGAGAGCTGTTAAAGAGCTGCGCTGAATTG 256
Qy 377 ATGAAGAGGTGCATGCACTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTAAATAG 436
Db 257 ATGCAAGAGGGCGCACTCTTTTAAACCAAGATGATGTGATGGATGGCGTGGCAAGCATG 316
Qy 437 GTTCCCGATCTAGGTGTAGAACCCACTTTCTGATGTGTACGAACCTTGTAACTGTGAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTCTGATGGGACTTAAACTCGTAACCGGTGAT 376
Qy 497 TGGCCCATCGAACCAAGATGAGCACTTCAAGCGGCGAAGTGAATTTGGTTGCGATAAA 556
Db 377 ACCCTATTGAGGCAATGGTAATTTAGTTCTGTTGAGTTG---TTCTTAAAAAATGNA 433
Qy 557 GACATCGAGCTCAATGCAGGCAAGAAAGTAAACCGAAGCTTGAGGTTACTTAATGAAGGCT 616
Db 434 GACATCACTATCAACGAAGGCAAGAAAGCGCTTAGCGTGAAGTTTAAAGATTTTGGCGAC 493
Qy 617 AAATCTTGTGATGGGTAGCCATTTCACCTTTCTTTGAAGCTAACAGGCACTAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCCCTAGACTTT 553
Qy 677 GATCGTAAAAAGCTATGCAAAACCGCTAGATATTCCTCTGCGCAACGCTACGCAATT 736
Db 554 GACAGAGAAAAAATTTTCGGTAAACCGTTAGACATTTGACGCGGAGCAGCGGTAAAGATT 613
Qy 737 GGGGACAGCAAAACCCGCAAGTGCAGTTGATTTCTTTGGTGGCAGTAAAGAAAGTGAAT 796
Db 614 GAGCTGCGGAGNAAATCGTGAATTTGATTGACATTTGCGGTAAACAGAGATCTTT 673
Qy 797 GGCATGAACGGGCTTGTGAATTAACATCGCGATGAACGCCATAAACAATAAGCGCTTGAC 856
Db 674 GGATTTAACGCTGGTTGATAGACAAGCAGCAACGAAGAAACAAAAAATTTGCTTTTACAC 733

PN W09109049-A.
 XX 27-JUN-1991.
 XX 19-DEC-1989; 89GB-00028625.
 XX 19-DEC-1989; 89GB-00028625.
 XX (THRE-) 3I RES EXPL LTD.
 XX Tabaqchal IS, Clayton CL, Wren BW, Kleanthous H;
 XX WPI; 1991-208084/28.
 DR P-PSDB; AAR12515, AAR13550.
 XX
 PT Oligo:nucleotide(s) specific for *Helicobacter pylori* - used as probes and
 PT primers to detect *H. pylori* infection, in diagnosis of gastritis, and
 PT duodenal and peptic ulceration.
 XX
 PS Disclosure; Page 16; 28pp; English.
 XX
 CC The DNA is a 2.7 kb *taqI* fragment encoding the A and B subunits of *H.*
 CC *pylori* (previously *C. pylori*) urease, i.e. the 66 and 31 kD antigens.
 CC From the sequence, probes and primers can be designed for the
 CC amplification (by PCR) of the gene, to produce a prods. common to all *H.*
 CC *pylori* strains so far tested and which do not occur in other ureases and
 CC can therefore be used as a specific indication of the presence of *H.*
 CC *pylori*. *H. pylori* (previously *C. pylori*) is strongly implicated in the
 CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
 CC primers/probes can be used for the detection of *H. pylori* DNA in gastric
 CC mucosa, saliva or faecal samples to provide an early diagnosis of
 CC infection. See also NAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
 CC OS field.)
 XX
 SQ Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;
 Query Match 36.5%; Score 1053.6; DB 2; Length 2767;
 Best Local Similarity 65.7%; Pred. No. 2.5e-226;
 Matches 1615; Conservative 0; Mismatches 804; Indels 39; Gaps 4;
 182 AAAATTAAACACAAAGAGTAATAGGTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTG 241
 40 ACACCTTAAAGATPAGGAGATGAGTGAATCACTCACCCCAAGAGTTAGACAAAGTTGATG 99
 242 TTATATATGCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAA 301
 100 CTCCTACTACGCTGGAGAAATTTAGCTAATAAAACGCAAGAAAGGCAATTAAGCTTAACTAT 159
 302 CCCGAAGCCATTGCTTACATTAGTGCCTCATATTATGACGAAGCGCGCGTGGAAAAAAA 361
 160 GTGGAAAGCGGTACGTTTGTATTGATGCCCATATTATGGAAGAGAGAGAGAGCTGGTAAAG 219
 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGGAAAAAGATGAAGTAAATGCC 421
 220 ACTGCGGCTGAATTGTATGATGAAGAGGCGGCGCACTCTTTTAAACCGGATGATGTGGAT 279
 422 GGGGTGGGTAAATATGGTTCCTGATCTAGGTGTAGAGCCACCTTCTCTGATGTAGAA 481
 280 GCGGTGGCAAGCATGATCCATGAATGGGTGATTTGAAGCCGATGTTCTGATGGGACCAA 339
 482 CTTGTAACTGTGAATTTGGCCATCGAACACAGATGAGCATTCAAAGCGGCGAAGTGAAA 541
 340 CTCGTAAACCGTGCATACCCCTATTGAGGCCAATGGTAAATATTGTTCTCGGTGAGTTG--- 396
 542 TTTGGTTGGCTAAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAAACCGAACTTTGAGGTT 601
 397 TTCTTAAAAAATGAAGACATCACTATCAAGAGGCAAAAAGCCGTAGCGTGAAGTT 456
 602 ACTAATGAAGGGCTTAAATCTTTGCAATGGGTAGCCATTTCACTTCTTTTGAAGCTAAC 661
 457 CCCCCTGTTGGGACAGACGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 516
 662 AAGGCACTAAATTCGATCGTGAAGAGCCATATGGCAAGCCCTAGATATTCCCTCTGGC 721

Db AGATGCTTAGACTTTTGAACAGAGAAAAAACTTTCGGTAAACCGCTTAGACATTCGAGCGGG 517
 Qy AACACGCTTACGATTTGGGCGAGGACAAACCCGCAAAAGTGCAGTTGATTCTCTTGGTGGC 781
 Db ACAGCGGTAAAGTTTGAAGCTCGCGCAAGAAAAATCCGTAGAGTTGATTTGACATTTGGCGGT 636
 Qy AGTAAAAAAGTGAATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAAA 841
 Db AACAGAAAGATCTTTTGGATTTAACCGGTTGGTTGATAGGCAAGCCGATAAACGAAAGCAAA 696
 Qy CATAAAGCGCTTGACAAGGCGAAATCTCACGGATTT----- 877
 Db AAAATTGCTTTTACACAGAGCTTAAAGAGCGTGGTTTTCATGGCGCTTAAAGCGATGACAAC 756
 Qy -----ATCAAGTAAAGGAGACTCCCATGAAA---ATGMAAAAAACAGAAATATGTAAAT 926
 Db TATGTAAAAACATTTAAGGAGTAAAGAAATGAAGAATAGCAGAAAGAAATATGCTTCT 816
 Qy ACCTACGGAACCCCAAGGCGGATAAAGTGGCTTTAGGAGATACCGATCTTTTGGGCGAGAA 986
 Db ATGTATGGCTTACTACAGGCGATTAAGTGAGATTGGGCGATACAGACTTTGATCGCTGAA 876
 Qy GTAGAACATGACTATACACCTATGCGGAGAACTTAAATTTGGCGCGGTAAAACTATC 1046
 Db GTAGAACATGACTATACACCTATGCGGAGAGCTTAAATTTGGCGCGGTAAAAACCTTA 936
 Qy CGTGAGGATGCTGCTGAGGCAATAGCCCTGATGAACACACCTAGATTTAGTCACTACT 1106
 Db AGAAGGCGATGAGCCCAATCTAACACCTTAGCAAGAAAGAACTGGATCTAATCATCACT 996
 Qy AACCGCATGATTTACGCACTACACCGGATTTACAAAGCGGACATTTGGGATTTAAAAACGCG 1166
 Db AACGCTTTAATCGTGGATTACCGGATTTTAAAGCGGATTTGTTATTTAAGATGGC 1056
 Qy AAAATCCATGGCATTTGGCAAGGCGAGAAACAAAGGACATGCAAGATGGCGTAAAGCCCTCAT 1226
 Db AAAATCGCTGGCATTTGGTAAAGGCGGTAAACAAAGACACGCAAGATGGCGTTAAAAACAAT 1116
 Qy ATGTCGTGGGTGGGCGACAGAGCACTAGCAGGGAAGGTATGATTTATTACCGCTGG 1286
 Db CTTAGCGTGGGTCTCTGCTACTGAAAGCCCTTAGCCGTTGAAGTTGATTTGTAATCGTGGT 1176
 Qy GGAATCCATTCACACACCCCTCTCTTCTCCACAAACAATTCCTACCGCTCTAGGCAAT 1346
 Db GGTTATGACACACACATCCCTCTCTCTCCCAACAAATTCCTTACAGCTTTTGAAGC 1236
 Qy GGCTTTACAAACCATGTTTGGAGCGGACACAGGTCTCTGTAGATGGCACGAATCGCACTACT 1406
 Db GGTGTAAACCATGATTTGGTGGCGGAACCTGGCCCTGCTGATGGCACTTAAACGCAACCACT 1296
 Qy ATCACTCCGGGCAATGGAACTTGCACCGCATGTTTGGCGGACAGAGAGATTTCTATG 1466
 Db ATCACTCCAGGTAGAGAAATTTAAAAATTCATGTCTCAGAGCGGCTTGAAGAAATTTCTATG 1356
 Qy AATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAACAACTTTGTAGAACAGTA 1526
 Db AACTTTGGTTCTTTGGCTAAAGGTAAACGCTTCTAACGATGCAAGCTTAGCCGATCAAAAT 1416
 Qy GAAGCGGCGCGATTTGGTTTAAATTTGATGAAGACTGGGCGCAACCAACCAAGTGCATC 1586
 Db GAAGCTGGTGGATTTGGCTTTAAATCCAGAGAGCTGGGCGACCACTCTCTTCTGCAATC 1476
 Qy GATCACTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACA 1646
 Db AATCATCGTTAGATTTGCGGACAAATACGATGTGCAAGTTCGCTATCCACACAGACACT 1536
 Qy GTCAATGAGGCGGTTATGATGACACCCCTAAATGCAATGAAGAGGCGGCGCCATCCAT 1706
 Db TTGAATGAAGCGGTTGCGTGGAAAGACACTATGCGAGCTATTGCGGAGCGACATATGAC 1596
 Qy GCCTTACCACATTTAGGAGCGGTTGGAGGACACTCACTGATGTTTATCACCATGGCAGGC 1766

Db 1597 ACTTACCACACTGAAGCGCTGGCGGCGACAGCTCCTGTGATATTATTAAAGTGGCGGT 1656
Qy 1767 GAGCTCAATATTCTACCTCTCCACACCCACCCACTATTCCTATACCAATTAATACGGTT 1826
Db 1657 GAAACAACATCTCCCGCTTCCACTAACCCCACTATCCCTTCACCGTGAATACAGAA 1716
Qy 1827 GCAGAACACTTAGACATGCTCATGACATGCCACCACTTAGACAAACGCAATCCGCGAGGAT 1886
Db 1717 GCCGAACACATGGACATGCTTATGGTGTGCCCACTTGGATATAAAGCATTAAGAAGAT 1776
Qy 1887 TTACAAATTTCTCAAGCGCTATCGCGCGCTCTATCGCGCTGAAGATGTGCTCCAT 1946
Db 1777 GTCCAGTTCGCTGATTCAAGGATTGCCCTCAAAACCAITTCGCGCTGAAGCACTTTGCAT 1836
Qy 1947 GATATGGGTGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGGGTGCAGCGAA 2006
Db 1837 GACATGGGATTTTCTCAATCACTAGTCTGACTCTCAAGCGATGGCGGTGTGGGTGAA 1896
Qy 2007 GTGATTCCTCGAACTTGGCGAGACTGCGGATAGATATAAAGAAATTTGGTAACTTCCT 2066
Db 1897 GTTATCACTAGAACTTGGCAACAGCTGACAAAAATAAAAAAGAAATTTGGCGCTTGAAA 1956
Qy 2067 GAAGATGGCAAGATAACGATAATTTCCGCAATTAAGCGCTACATCTCCAATACACTATC 2126
Db 1957 GAAGAAAAAGCGGATAACGACAACTTCAGGATCAAAACGCTACTGTCTAAATACACATT 2016
Qy 2127 AACCCCGCTTTGACCCACCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGCGAAGATC 2186
Db 2017 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTCGGTCTGTAGAAAGTGGCAAGTG 2076
Qy 2187 GCCGACTTGGTGGTGGGAATCTGCTTTTGGCGTAAAAACCCCAAAATCGTGATCAAA 2246
Db 2077 GCTGACTTGGTATTGTGGAGTCCGCAATTCCTTGGTGTGAAACCCCAACATGATCATCAA 2136
Qy 2247 GCGGTATGGTGGTCTTCTCTGAATGGCGGATCTTAACGCTCTGTGCCACTCCCA 2306
Db 2137 GCGGGTTCATCGCATTGATGATCAATGAAATGGGTGATGCGAACGCTTCTATCCCTACCCACAA 2196
Qy 2307 CCGGTTTATTACCGCAAAATGTTGGGCATCAGCGCAAGCGCAAAATTTGACACCCAGCATC 2366
Db 2197 CAGTTTATTACAGAGAAATGTCGTCATCATGTGAAGCTTAATACGATGCAACATC 2256
Qy 2367 ACTTTTGTTCAAAGTCCGCTATGAAATGGCGTGAAGAAAGCTGGCTTAGAGCGC 2426
Db 2257 ACTTTTGTCTCAAGCGCTTATGACAAAGCATTAAGAAAGAAATTAAGGCTTGAAGA 2316
Qy 2427 CAAGTCTACCGGTCAAAAACCTGCCGTAACTACATCACCAGAAAGACTTCAAGTTCAACGAC 2486
Db 2317 CAAAGTGTTCGCGTAAAAAATTTGCAGAAATATCACTAAAAAAGACATGCAATTTCAACGAC 2376
Qy 2487 AAAACGGCAAAATCACCGTTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAACTC 2546
Db 2377 ACTACCGCTCAGATTGAAGTCAATTCGAAACTTACCATGTGTCTGTGATGGCAAA-GA 2435
Qy 2547 TGCACCTCTAAACCCCACTCGCAAGTGCCTTAGCCCCAGCGGTACACTTCTTCTTAGG 2604
Db 2436 AGTAACTTAACACCGCAATAAAGTGAGCTGGCGCAACTCTTAGCATTTTCTAGG 2493

RESULT 14

ADQ37850
ID ADQ37850 standard; DNA; 1710 BP.
XX AC ADQ37850;
XX XX
DT 07-OCT-2004 (first entry)
DE H. bizzozeronii ureB DNA.
XX Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW antibacterial; gene; ds.
XX

OS Helicobacter bizzozeronii.
XX
FH Key Location/Qualifiers
FT 1. .1710
FT CDS /*tag= a
FT /product= "H. bizzozeronii ureB"
XX
PN US2004142343-A1.
XX 22-JUL-2004.
XX
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHU/) ZHU J.
XX
XX Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
XX P-PSDB; ADQ37851.
XX
XX Novel isolated nucleic acid molecule having urease gene cluster, and
XX conferring on Helicobacter bizzozeronii ability to produce urease, useful
XX as vaccine for preventing disease in mammals infected by H.bizzozeronii.
XX
XX Claim 7; SEQ ID NO 4; 40pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule conferring on
XX Helicobacter bizzozeronii an ability to produce urease, where the nucleic
XX acid molecule is a urease gene cluster comprising at least one urease
XX structural gene and at least one urease accessory gene. The nucleic acid
XX molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
XX invention also relates to an isolated protein encoded by the nucleic
XX acid, a vaccine for preventing onset of disease in mammals infected by H.
XX bizzozeronii comprising a nucleic acid and a carrier, and an isolated
XX antibody or its binding portion raised against the nucleic acid. The
XX nucleic acids, proteins and antibodies are useful for vaccinating mammals
XX against onset of disease caused by infection of H. bizzozeronii, which
XX involves administering the sequences. The sequences are useful for
XX detecting H. bizzozeronii in a sample of tissue or body fluids which
XX or providing a nucleic acid as an antigen, providing an antibody,
XX or providing a nucleotide sequence as a probe in a nucleic acid
XX hybridisation assay, contacting the sample with the antigen or the probe,
XX and detecting any reaction which indicates that H. bizzozeronii is
XX present in the sample. This sequence represents H. bizzozeronii ureB DNA.
XX
XX Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;
XX
Query Match 32.6%; Score 940.8; DB 12; Length 1710;
Best Local Similarity 72.0%; Pred. No. 4.8e-201;
Matches 1227; Conservative 0; Mismatches 477; Indels 0; Gaps 0;
Qy 900 AAAATGAAAAACAGAAATATGTAATACCTACGACCCACCAAGGCGATAAAGTGGC 959
Db 7 AAAATCTCTGAAAAGAAATATGTTCTATGTATGACCCACTACGGCGATAAAGTGAGA 66
Qy 960 TTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGGCGAAGAA 1019
Db 67 TTGGCGATACCGACCTGATCTTTAGAGTGAACATGACTGCACCCTTATGGCGAAGAA 126
Qy 1020 CTTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATGGTTCAGAGCAATAGCCCTGAT 1079
Db 127 ATTAAGTTTGGTGGCGTAAAAACCATTCGCGATGGGATGGCAAAACCAACAGCCCCAGC 186
Qy 1080 GAAAAACCTAGATTTAGTCATCACTAACGCGATGATTATCGACTACACCGGATTTAC 1139
Db 187 AGCCAGAACTCGATCTTGTGCTCACTAACGCCCTGATCGTGGATTACACCGCATTTAT 246
Qy 1140 AAAGCCGACATTGGGATTAAAAACGGCAAAATCCATGGCATTTGGCAAGCGCAGGAACAAG 1199

Db 247 AAGCCGATATTGGCAATTAATAATGGCAAAATCCATGGCAATTGGCAAAAGCAGGCATATAA 306
Qy 1200 GACATGCAAGATGGCGTAAGCCCTCATATGGTGGTGGTGGGCAACAGAACCACTAGCA 1259
Db 307 GACATGCAAGATGGCGTTTGCAACATCTTTGGTGGGCCCTGCTACTGAGGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATATTACCGTGGGGGAATCGAATTCACACCCACTTCTCTTCTCCA 1319
Db 367 GCTGAAGGGCTGATTGTTACAGCTGGTGGATTGACACCCACATCCACTTTATTCTCTCC 426
Qy 1320 CACAAATTCCTTACCGCTTAGCCATGGCTTACACCAATGTTTGGAGGGGACACAGGT 1379
Db 427 CAACAAATCCCAAGCATTTGCGAGCGGATCAACACCAATGATTGGTGGGGGAACAGGT 486
Qy 1380 CTGTAGATGGCAGAAATGGACTACTATCACTCCGGGCAATGGAATCTTGACCGCATG 1439
Db 487 CAGCTGATGGGACTAAGCGGACTACATCACTCCGGGGCGCTGGAACCTTTAAACCATG 546
Qy 1440 TTGCGCGCAGCAGAAGATATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGGTGCCTCTGAAGAATATGCCATGAACTTGGGCTATTGGGTAAAGGGAATGTCT 606
Qy 1500 AGCAAAACAACTTTAGAACAAAGTAGAAGCGGGCCGATTTGGTTTAAATTGCATGAA 1559
Db 607 TATGAACCCCTCCCTGGTTCGATCAACTCGAAGCTGGAGCCATTGGCTTTAAATCCAGAA 666
Qy 1560 GACTGGGGCACACACCAAGTGGATCGATCACTGCTTGGCTGGCGGAGATGAATGATGAT 1619
Db 667 GACTGGGGTAGCACACCTCGAGGCATCTACCATTTGCTTGAATGTGGCTGCAAAATACGAT 726
Qy 1620 GTCAAGTTTGTATCCACACCCGATACAGTCAATGAGCGAGGTATGTAGATGACACCCTA 1679
Db 727 GTGCAAGTGGCTATCCACCCGATACCTTGAATGAAGCGGGCTGTGTGAAGACACTTGTG 786
Qy 1680 AATGCAATGAACGGGCGCGCCATCCATCCCTACACATTCAGGAGAGCGGTGGAGGACAC 1739
Db 787 CAAGCCATTGCTGGGCGCATATCCACACTTCCACACTGAAGGTGCTGGTGGGGGCAC 846
Qy 1740 TCACCTGATGTTATCACCATGGCAGGAGCTCAATATTCTACCTCTCTCCACCACCC 1799
Db 847 GCTCCGGATGTCATTAAGATGTCTGGCGAATTTAACATCTCTCCAGCTTCTACCAACCC 906
Qy 1800 ACTATTTCCCTATACCATTAATACGGTTGACAGACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACATTTCTTTTACCGTGAATACAGAACCCGACACATGACATGTTGATGGTGGCCAC 966
Qy 1860 CACTAGACAAACGCTATCCGCGAGGATTTA CAATTTTCTCAAAGCCGTATCCGCCCGGC 1919
Db 967 CACTTGGATAAACAATCAAGAAGATGTCCAGTTTGTCTGATTCTAGGATTGGCCCCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGTCTCATGATATGGGTGTGATTCGGGATGACACAGCTCGGAT 1979
Db 1027 ACCATCGCGCTGAGGACAACTCCACGATATGGGATTTTCTCTATCACCAGCTCTGAC 1086
Qy 1980 TCGCAAGCAATGGGCGTGCAGCGGAGTATCTCTGCACTTGGCAGATGCGGATAAG 2039
Db 1087 TCCCAAGCAGATGGGCGTGTAGCGAGGTCAATCACCCTCTTGGCAACAGCGGCAAA 1146
Qy 2040 AATAAAAAGAAATTTGGTAGCTTCTCTGAAGATGGCAAGATAACGATAATTTCCGCAAT 2099
Db 1147 AACAAAAAGAAATTTGGTGGCTTGGCTGAGGAAAGGCGGATATGACAACTTCGGATC 1206
Qy 2100 AAGCGCTACATTTCCAAATACATATCAACCCCGCTTTGACCCACGGCGTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTGACACAGGCAATTTCTGAATAT 1266
Qy 2160 ATCGGCTGTGGAAGGCAAGATCGCCGACTTGGTGTGTGGAATCTTGCCTTTTTT 2219
Db 1267 GTCGGCTCTGTAGAAGTGGGCAAAATTCGCCGATTTGGTGTCTTGGAGTCTCTGGCTCTTT 1326
Qy 2220 GGGGTAAACCCAAATCGTGATCAAGGCGGATGTTGTTCTCTCTGAAATGGGCGAT 2279
Db 1327 GGCATTAACCCCAACATGATCATCAAGGCGGATTCATCGCACTTTCTCAATGGGCGAT 1386

Qy 2280 TCTAAACGGTCTGTGCCACTCCCAACCCGGTTTATTACCGGAAATGTTTGGGCAATCAC 2339
Db 1387 GCCAATGCTCTATCCCACTCCCAACCCCGTATTATTACCGGAAATGTTTGGCCCACT 1446
Qy 2340 GGCAGGCGAAATTTGACACCAGCATCACTTTTGTTCCTCAAGTCGCTTATGAAATGCG 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCACTTTTGTATCCCAAGTGGCTTATGACAAAG 1506
Qy 2400 GTGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTTCTTACCGTCAAAAACCTGCGGTAAACATC 2459
Db 1507 ATTAAGAGAGTTGGGCTTCAAGAGTGGTTTGGCAGTTAAACCTGCGGCAACATC 1566
Qy 2460 ACCAAGAAAGCTTTCAAGTTTCAACGACAAAAACGCGAAAAATCACCGTCAATCCGAAAAAC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCACGATTTACCGCACATCGAAGTCAATCTCTGAAAC 1626
Qy 2520 TTCGAGTCTTTGTAGATGGCAAACTTCGACCTCTAAACCCACTCGCAAGTGCCTCTA 2579
Db 1627 TACAAAGTTAAAGTGGATGGCAAGAGGTTTCTTCCAAAGCAGCGGATAAAATCAGCCTA 1686
Qy 2580 GCCAGCGCTACACTTTCTTCTAG 2603
Db 1687 GCACACTCTACAACTTGTCTAG 1710

RESULT 15

ABA00816
ID ABA00816 standard; cDNA; 1719 BP.
XX ABA00816;
AC ABA00816;
DT 01-APR-2003 (first entry)
XX H. felis urease B coding sequence.
DE Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
XX immune response; circulatory vessel; Gut Associated Lymphoid Tissue;
KW GALT; immune response; antigen; Salmonella; Cholera; immunity;
KW Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
XX hormone; hormone receptor; cancer; ss.
OS Helicobacter felis.
XX WO200296949-A1.
PN 05-DEC-2002.
XX 24-MAY-2002; 2002WO-AU000661.
PD 25-MAY-2001; 2001AU-00005241.
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA Mckenzie B, Boyle J, Lew A;
XX WPI; 2003-156789/15.

Raising an immune response in an animal by administering composition comprising carrier and antigen bound to targeting moiety which binds to receptor present in circulatory vessels in Gut Associated Lymphoid tissue.

Disclosure; Page 27-28; 45pp; English.

This sequence encodes H. felis urease B. This sequence was used in the method of the invention for raising an immune response in an animal. The method comprises administering to the animal a composition comprising a carrier and an antigen bound to a targeting moiety which binds to at least one receptor present in circulatory vessels in Gut Associated Lymphoid tissue (GALT). The method is useful for raising an immune response in an animal against antigens from Salmonella, Cholera, Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut

CC	associated toxins, gut hormones, gut hormone receptors or gut associated
CC	cancers. The method is useful for raising both mucosal and systemic
CC	immunity against any antigen used in the composition
XX	
SQ	Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;
	Query Match 31.3%; Score 902.4; DB 8; Length 1719;
	Best Local Similarity 70.6%; Pred. No. 2e-192;
	Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;
QY	900 AAAATGAAACACAGAAATATGTAATACCTAGGACCCACCAAGGCGATAAAGTGGCG 959
DB	15 AAGATTTCAGAAAGAAATATGTTCTATGTATGTFCCCACTACCGGGGATCGTGTAGA 74
QY	960 TTAGGAGATACCGATCTTTGGGCGAAGTAGAACATGACTATACCACTATGCGCAAGAA 1019
DB	75 CTCGGGACACTGATTTGATCTTAGAAGTGGAGCATGATTGCACCACTTATGTTGAAGAG 134
QY	1020 CTTAAATTTGGCGCGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
DB	135 ATCAAAATTTGGGGCGGTAAAACTATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGC 194
QY	1080 GAAACACCTAGATTTAGTCACTCACTAACCGCATGATTATCGACTACACCGGATTTAC 1139
DB	195 TCTTATGAATTAGATTTGTTGCTCACTAACCGCCTCAITTTGGACTATACGGGCATTTAC 254
QY	1140 AAAGCCGACATTTGGGATTAATAACGGCAAAATCCATGGGATTTGGCAAGGCGAGAAACAAG 1199
DB	255 AAAGCCGACATTTGGGATTAAGAAGCGCAAGATTGAGGCATTTGGCAAGGCGAGCAATAG 314
QY	1200 GACATGCAAGATGGCGTAAAGCCCTCATATGTTGTTGGGTGGGCAAGAGCACTAGCA 1259
DB	315 GACATGCAAGATGGCGTAGATAATAATCTTTGGGTAGGTCCTGCTACAGAGGCTTTGGCA 374
QY	1260 GGGGAAGGTATGATTTATACCGCTGGGGAATCGATTACACACCACTTCCTTTCTCCA 1319
DB	375 GCTGAGGGCTTGATTTGTAACCGCTGGTGCATCGATACGCATATTCACATTTATCTCTCC 434
QY	1320 CAACAAATCCCTACCGCTCTAGCCATGGGTTTACAAACCATGTTTGGAGCGGCACAGGT 1379
DB	435 CAACAAATCCCTACCTTTTGGCCAGCGGGTTTACAAACCATGATTGGAGGAGCACAGGA 494
QY	1380 CTTGTAGATGGCAAGAAATCGCACTATCATCTCCGGGCAAAATGGAACTTGCAACCGCATG 1439
DB	495 CTTGCGGATGGCAAGAAATCGCACTATCATCTCCCGGCGCGCTTAATCTTAAAGATAG 554
QY	1440 TTGCGGCGAGCAGAGATTTCTATGATGTTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
DB	555 TTGCGTGCAGCGGAGAAATACGCCATGAATCTAGGCTTTTGGCTAAGGGGAATGTGTCT 614
QY	1500 AGCAAAACAACTTGTAGAACAGTAGAGCGGCGGATTTGTTTAAATTTGCATGAA 1559
DB	615 TAGCAACCTCTTTACGCAATCAGATTGAAGCAGGGCGGATTTGTTTAAATTTCCACGAA 674
QY	1560 GACTGGGCGACACCAAGTGGATCGATCACTGCTTTGAGCGTGGCAGATGAATACGAT 1619
DB	675 GACTGGGGAAGCACACTTCGAGCTATTCACCACTGCTCAATGTCGCGATGAATACGAT 734
QY	1620 GTGCAAGTTTGTATCCACACCGGATCAGTCAATGAGCGAGGTTATGATGATGACACCTTA 1679
DB	735 GTGCAAGTGGCTATCCACACCGGATCCCTTAAACGAGCGGGCTGTGTAGAAGACACCCCTA 794
QY	1680 AATGCAATGAAACGGGCGGCGCATCCATGCTTACCATTTGAGGGAGCGGTGGAGCAC 1739
DB	795 GAGGCGAATTTCCGGGCGGCAATCCATCCATACCTTCCACACTGAAGGGGCTGGGGTGGACAC 854
QY	1740 TCACCTGATGTTATACCAATGGCAGCGAGCTCAATATTTTACCTCTCCACACCCCC 1799
DB	855 GCTCCAGATGTTATCAAAATGGCAGGGGAATTTAACTTACCCGCTCTACTAACCCCG 914
QY	1800 ACTATTTCCCTATACCAATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCAC 1859
DB	915 ACCATTCTCTTACCAAAACACTGAAGCCGAGCACATGGACATGTTAATGGTGGCCAC 974

QY	1860 CACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTCTCAAAGCCGATCCGCCCGGC 1919
DB	975 CACTTGGATAAAAGATATCAAGGAAGATGTGCAGTTTGCAGATTCGAGGATTCGCCCCCAA 1034
QY	1920 TCTATCGCGGTGAAGATGTCTCCATGATATGGTGTGATCGGATGACAAAGTCCGAT 1979
DB	1035 ACTATCGCGCTGAAGACCAACTCCATGACATGGGATCTTTCTATACCAAGTCCGAC 1094
QY	1980 TCGCAAGCAATGGGGCTGAGCGGAAGTGAATCTCTCGAACTTGGCAGACTGCGGATAAG 2039
DB	1095 TCTCAGCTATGGAGCGGTAGGGAGGTGATACACGCACTTGGCAGACAGACAAA 1154
QY	2040 AATAAAAAAGAAATTTGGTAAAGTTCCTGAAAGTGGCAAGATAACGATAATTTCCGCAAT 2099
DB	1155 AACAAAAAGAGTTTGGGCGTTTGAAGAGAAAAAGCGGATTAACGACAACTTCGCGATC 1214
QY	2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACGGGTGAGGAGTAT 2159
DB	1215 AAACGCTACATCTCTAAATAACCATCAACCCCGCATCGCGCATGGGATTTCTGACTAT 1274
QY	2160 ATCGGCTCTGTGAAGAGGCAAGATCGCGACTTGGTGGTGTGGAATCTCTGCAATGGCGAT 2219
DB	1275 GTGGGCTCTGTGAAGTGGGCAAAATACCGCACTCTGCTTTGGAGTCCGCTTTCTTT 1334
QY	2220 GGCTTAAACCCAAATTCGTGATCAAGGGCGGTATGGTGGTCTTCTCTGAAATGGCGAT 2279
DB	1335 GGCAATTAAGCCCAATATGATTAATTAAGGGCGGATTTATTTGGCTCTCTCAATGGCGAT 1394
QY	2280 TCTAACGGCTCTGTGCCACTCCCAACCGGTATTTACCGCGAAATGTTTGGGATCAC 2339
DB	1395 GCCAATGCGTCTATTTCCACCCCTCAGCCGCTCTATTACCGTGAATGTTTGGACACCAT 1454
QY	2340 GGCAAGGCGAAATTTGACACCAAGCATCACTTTTGTTCBAAAGTCGCTATGAAATGCG 2399
DB	1455 GGGAAAAACAAATTCGACACCAATATCACTTTCTGTGTCGCAAGCGGCTTCAAGCGAGG 1514
QY	2400 GTCAAAAGAAAAGCTGGGCTTTAGAGCGCCAAAGTTCTACCGGTCAAAAACTGCCGTAAACATC 2459
DB	1515 ATCAAGAGAAGAACTAGGGCTAGATCGGTGATTTGCCAGTGAAGAACTGTGCAATATC 1574
QY	2460 ACCAAGAAAGACTTCAAGTTCAACGACAAACGGCAAAAAATCACCGTCGATCCGAAAAACC 2519
DB	1575 ACTAAAAAGGACCTCAAAATTCACGATGTGACCGCACATATTGATGTCAACCTGAAAAACC 1634
QY	2520 TTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTTAAACCCACCTCGCAAGTGCCTTA 2579
DB	1635 TATAAGGTGAAGTGGATGGCAAGAGGTAAACCTCTAAAGCAGCAGATGAATTTGAGCCTA 1694
QY	2580 GCCAGCGCTACACTTTTCTTCTAG 2603
DB	1695 GCGCACTTTTATTAATTTGTTCTAG 1718

Search completed: November 29, 2005, 00:34:36
Job time : 1579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:08:23 ; Search time 2049 Seconds
(without alignments)

11635.245 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 2883

Sequence: 1 tggagatttccarctt.....aaaaagtagagccacag 2883

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134699005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2880.6	99.9	2883	3	US-09-904-994B-1
2	2228	77.3	2452	3	US-09-904-994B-13
3	2158.6	74.9	2405	3	US-09-904-994B-4
4	2158.6	74.9	2407	3	US-09-904-994B-10
5	1948	67.6	2183	3	US-09-904-994B-7
6	1184.4	41.1	8407	7	US-10-639-273-1
7	940.8	32.6	1710	7	US-10-639-273-4
8	902.4	31.3	1719	8	US-10-476-313-10
9	886.4	30.7	1710	7	US-10-335-977-3849
10	885.8	30.7	1815	3	US-09-895-913A-251
11	884.8	30.7	1710	7	US-10-282-122A-22427
12	880	30.5	1710	9	US-10-500-447A-5
13	876.8	30.4	1717	8	US-10-476-313-11
14	855.6	29.7	2385	3	US-09-402-100-1
15	734.6	25.5	2341	7	US-10-282-122A-32950
16	690.6	24.0	2439	7	US-10-282-122A-34390
17	664.4	23.0	1830121	7	US-10-329-670-1
18	664.4	23.0	1830121	8	US-10-158-865-1
19	664.4	23.0	1830121	9	US-10-981-687-1
20	662.8	23.0	5966	2	US-08-781-986A-22
21	662.8	23.0	5966	7	US-10-329-624-22
22	658.6	22.8	1698	7	US-10-282-122A-33102
23	634.8	22.0	3919	9	US-10-795-159-460

ALIGNMENTS

RESULT 1

US-09-904-994B-1
; Sequence 1, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1e1 N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degosedesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2883
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (206)..(886)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (897)..(2603)
US-09-904-994B-1

Query Match 99.9%; Score 2880.6; DB 3; Length 2883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	RGGRAGATTTCCARCACTTCACGACATATTTGATCTGTGGTGGTAAATTCR	60
Db	1	RGGRAGATTTCCARCACTTCACGACATATTTGATCTGTGGTGGTAAATTCR	60
Qy	61	ACTTGTTAACTCTATTATTAATTTTAAATTAATTTACTTATTATCATATATAATATTA	120
Db	61	ACTTGTTAACTCTATTATTAATTTTAAATTAATTTACTTATTATCATATATAATATTA	120
Qy	121	TTACTTATTAATAAGTTAATAAAAAAGTAAACGAAATTTAGACTATAATCCCATGGCCTT	180
Db	121	TTACTTATTAATAAGTTAATAAAAAAGTAAACGAAATTTAGACTATAATCCCATGGCCTT	180
Qy	181	TAAATTTAACACAGAGTAATAGGTGAACACTCACACCCAAAGCAGCAAGAAAGTCTT	240
Db	181	TAAATTTAACACAGAGTAATAGGTGAACACTCACACCCAAAGCAGCAAGAAAGTCTT	240
Qy	241	GTTATATTATGCGGGCGAAGTGGCTAGAAACGCAAGCAGAGGGCTTTAAAGCTCAACCA	300

Db	241	GTTATATTATGCGGCGGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCA	300
Qy	301	ACCCGAAGCCATTGCTTACATTAGTCCCATATTATGGAAGAGCGCGCGTGGAAAAAA	360
Db	301	ACCCGAAGCCATTGCTTACATTAGTCCCATATTATGGAAGAGCGCGCGTGGAAAAAA	360
Qy	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCATCTTTTGGAAAAAGATGAAGTAATGCC	420
Db	361	AACCGTTGCCAGCTTATGGAAGAGTGCATGCATCTTTTGGAAAAAGATGAAGTAATGCC	420
Qy	421	CGGGTGGGTAAATAGGTTCCCGATCTAGGTGTAGAGCCACTTTCCTGATGGTACGAA	480
Db	421	CGGGTGGGTAAATAGGTTCCCGATCTAGGTGTAGAGCCACTTTCCTGATGGTACGAA	480
Qy	481	ACTTGTAATCTGTAATTTGGCCCATCGAACCAAGATGACACTTCAAGCGGGGGAAGTAA	540
Db	481	ACTTGTAATCTGTAATTTGGCCCATCGAACCAAGATGACACTTCAAGCGGGGGAAGTAA	540
Qy	541	ATTGGTTGCGATAAAGACATCGAGCTCAATGCAAGCAAAAGATGAACCGAACTTGAGGT	600
Db	541	ATTGGTTGCGATAAAGACATCGAGCTCAATGCAAGCAAAAGATGAACCGAACTTGAGGT	600
Qy	601	TACTAATGAAGGCTTAATCTTGCATGTGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
Db	601	TACTAATGAAGGCTTAATCTTGCATGTGGTAGCCATTTCCACTTCTTTGAAGCTAA	660
Qy	661	CAAGCACATAAATTCGATCGTGAAGGCTATGGCAAGCGCTAGATATTCCCTCTGG	720
Db	661	CAAGCACATAAATTCGATCGTGAAGGCTATGGCAAGCGCTAGATATTCCCTCTGG	720
Qy	721	CAACACGCTACGCAATTTGGGCGAGGCAAAACCCGCAAAAGTGCAGTTGATTCCTCTGGTGG	780
Db	721	CAACACGCTACGCAATTTGGGCGAGGCAAAACCCGCAAAAGTGCAGTTGATTCCTCTGGTGG	780
Qy	781	CAGTAAAAAGATGATGGCATGAAACCGGCTTGTGAATTAATCATCGCGGATGAACGCCATAA	840
Db	781	CAGTAAAAAGATGATGGCATGAAACCGGCTTGTGAATTAATCATCGCGGATGAACGCCATAA	840
Qy	841	ACATAAGGCTTTGCAAGCGCAAAATCTCAGCGGATTTAACAAGTAAGGAGACTCCCATGA	900
Db	841	ACATAAGGCTTTGCAAGCGCAAAATCTCAGCGGATTTAACAAGTAAGGAGACTCCCATGA	900
Qy	901	AAATGAAAAAACAAGAATATGTAATACTTAACCGAACCAAGCGGATAAAGTGGCT	960
Db	901	AAATGAAAAAACAAGAATATGTAATACTTAACCGAACCAAGCGGATAAAGTGGCT	960
Qy	961	TAGAGATACCGATCTTTGGGAGAGATAGAAATGACTATACCACTATGGCGAAGAAC	1020
Db	961	TAGAGATACCGATCTTTGGGAGAGATAGAAATGACTATACCACTATGGCGAAGAAC	1020
Qy	1021	TTAAATTTGGCGGGTAAACTATCCGTGAGGATGGGTAGAGCAATAGCCCTGATG	1080
Db	1021	TTAAATTTGGCGGGTAAACTATCCGTGAGGATGGGTAGAGCAATAGCCCTGATG	1080
Qy	1081	AAAAACCCCTAGATTAGTACATCACTAAACCGCATGATATCGACTACACCGGATTTTACA	1140
Db	1081	AAAAACCCCTAGATTAGTACATCACTAAACCGCATGATATCGACTACACCGGATTTTACA	1140
Qy	1141	AAGCCGACATTTGGATTTAAAAACCGCAAAATCCATGGCATTTGCCAAGCAAAAGG	1200
Db	1141	AAGCCGACATTTGGATTTAAAAACCGCAAAATCCATGGCATTTGCCAAGCAAAAGG	1200
Qy	1201	ACATGCMAGATGGCGTAAGCCCTCATATGGTCTGGGTGGGCAAGAGCACTAGCAG	1260
Db	1201	ACATGCMAGATGGCGTAAGCCCTCATATGGTCTGGGTGGGCAAGAGCACTAGCAG	1260
Qy	1261	GGGAAGGTATGATTTATTCGCTGGGGGAATCGATTACACACCCACTTCTTTCTCCAC	1320
Db	1261	GGGAAGGTATGATTTATTCGCTGGGGGAATCGATTACACACCCACTTCTTTCTCCAC	1320
Qy	1321	AACAAATCCCTACCGCTTAGCCOATGGCGTTTCAACCAATGTTGGAGCGGCACAGGTC	1380
Db	1321	AACAAATCCCTACCGCTTAGCCOATGGCGTTTCAACCAATGTTGGAGCGGCACAGGTC	1380
Qy	1381	CTGTAGATGTCACGAATGCGACTTACTATCACTCGGGCAAAATGGAACCTTGCAACGCAATG	1440
Db	1381	CTGTAGATGTCACGAATGCGACTTACTATCACTCGGGCAAAATGGAACCTTGCAACGCAATG	1440
Qy	1441	TGCGCGCAGCAGAAGATATTCTATGAATGTGGCTTTTGGGCAAAAGCAATAGCTCTTA	1500
Db	1441	TGCGCGCAGCAGAAGATATTCTATGAATGTGGCTTTTGGGCAAAAGCAATAGCTCTTA	1500
Qy	1501	GCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGGATTTGGTTTAAATTTGCATGAAG	1560
Db	1501	GCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGGATTTGGTTTAAATTTGCATGAAG	1560
Qy	1561	ACTGGGGCAACAACCAAGTGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATG	1620
Db	1561	ACTGGGGCAACAACCAAGTGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATG	1620
Qy	1621	TGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCAAGTTATGTAGATGACACCCCTAA	1680
Db	1621	TGCAAGTTTGTATCCACCCGATACAGTCAATGAGGCAAGTTATGTAGATGACACCCCTAA	1680
Qy	1681	ATGCAATGAACGGCGCGCATCCATGCTTACCACTTGGAGGAGCGGTGAGGACACT	1740
Db	1681	ATGCAATGAACGGCGCGCATCCATGCTTACCACTTGGAGGAGCGGTGAGGACACT	1740
Qy	1741	CACCTGATGTTATCACCATGGCAGGAGCTCAATATTTCTACCCCTCTCCACCCCA	1800
Db	1741	CACCTGATGTTATCACCATGGCAGGAGCTCAATATTTCTACCCCTCTCCACCCCA	1800
Qy	1801	CTATTCCTTATACATTAATACGGTTGCGAGAACTTAGACATGCTCATGATGACACCC	1860
Db	1801	CTATTCCTTATACATTAATACGGTTGCGAGAACTTAGACATGCTCATGATGACACCC	1860
Qy	1861	ACCTAGACAAACCGCATCCGCGAGGATTAATAATTTCTCAAAGCGGTATCGCCCCGGCT	1920
Db	1861	ACCTAGACAAACCGCATCCGCGAGGATTAATAATTTCTCAAAGCGGTATCGCCCCGGCT	1920
Qy	1921	CTATCGCGGTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT	1980
Db	1921	CTATCGCGGTGAAGATGCTCCATGATATGGGTGATCGCGATGACAAGCTCGGAT	1980
Qy	1981	CGAAGCAATTTGGCGGTGCGAGGAGTGTCTCGAACTTTGGCAGACTGCGGATAAGA	2040
Db	1981	CGAAGCAATTTGGCGGTGCGAGGAGTGTCTCGAACTTTGGCAGACTGCGGATAAGA	2040
Qy	2041	ATAAAAAAGATTTGGTAAAGCTTCTGAAGTGGCAAGATAACGATATTTCCGATTA	2100
Db	2041	ATAAAAAAGATTTGGTAAAGCTTCTGAAGTGGCAAGATAACGATATTTCCGATTA	2100
Qy	2101	AGCGTACATCTCCAAATACATATCAACCCGCTTTGACCCACCGCGTGAAGGATTA	2160
Db	2101	AGCGTACATCTCCAAATACATATCAACCCGCTTTGACCCACCGCGTGAAGGATTA	2160
Qy	2161	TCGCTCTGTGGAAGGGCAAGATCGCCGACTTTGGTGTGGAACTCTGCTTTTGTG	2220
Db	2161	TCGCTCTGTGGAAGGGCAAGATCGCCGACTTTGGTGTGGAACTCTGCTTTTGTG	2220
Qy	2221	CGGTAAACCCCAATCTGATCAAGCGGTATGGTGTCTTCTGAAATGGGCGATT	2280
Db	2221	CGGTAAACCCCAATCTGATCAAGCGGTATGGTGTCTTCTGAAATGGGCGATT	2280
Qy	2281	CTAACCGTCTGTGCCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCACTACG	2340
Db	2281	CTAACCGTCTGTGCCCACTCCCAACCGGTTTATACCGGAAATGTTTGGGCACTACG	2340
Qy	2341	GCAAGCGAAATTTGACACCGCATCACTTTTGTGTTTCCAAAGTTCGCTATGAAAAATGGCG	2400
Db	2341	GCAAGCGAAATTTGACACCGCATCACTTTTGTGTTTCCAAAGTTCGCTATGAAAAATGGCG	2400
Qy	2401	TGAAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAAATGCGCGGTAAACATCA	2460
Db	2401	TGAAAGAAAAAGCTGGGCTTAGAGCGCAAGTTCACCGGTCAAAAAATGCGCGGTAAACATCA	2460

```
QY 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCAACGTCGATCCGAAACCT 2520
Db 2461 CCAAGAAAGACTTCAAGTTCAACGACAAACCGGCAAAAATCAACGTCGATCCGAAACCT 2520
QY 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCGCAAGTGCCTCTAG 2580
Db 2521 TCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCGCAAGTGCCTCTAG 2580
QY 2581 CCAGCGCTACACTTCTCTTAGGCAAAATGCCCCCTTTTGGGGGCGAGGTTATTTAGGAA 2640
Db 2581 CCAGCGCTACACTTCTCTTAGGCAAAATGCCCCCTTTTGGGGGCGAGGTTATTTAGGAA 2640
QY 2641 TCTTCATCAAAACGCACTGCAATCGGTCTTGGTGTGCGATGCTGCTTTAAAAACAAC 2700
Db 2641 TCTTCATCAAAACGCACTGCAATCGGTCTTGGTGTGCGATGCTGCTTTAAAAACAAC 2700
QY 2701 TTTTCATCTTTAGCAATCGCCATTTTAAATTAATTTAAATCTTATTAATTAATATAT 2760
Db 2701 TTTTCATCTTTAGCAATCGCCATTTTAAATTAATTTAAATCTTATTAATTAATATAT 2760
QY 2761 TATGCCCTCATTTTAAAGGAGAATTATGCGTAGGTCTTTGGTATTGCTATGTGGGG 2820
Db 2761 TATGCCCTCATTTTAAAGGAGAATTATGCGTAGGTCTTTGGTATTGCTATGTGGGG 2820
QY 2821 TTGTTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
Db 2821 TTGTTTGGTCTGGGCGCAAGGGTATTGAAACCCATCGCTCAAAAAAGTAGAAGCCAC 2880
QY 2881 AGG 2883
Db 2881 AGG 2883
```

RESULT 2

```
US-09-904-994b-13
; Sequence 13, Application US/09904994b
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1e1 N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degodequences
; CURRENT APPLICATION NUMBER: US/09/904, 994b
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 2452
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(728)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739)..(2445)
US-09-904-994b-13
```

```
Query Match 77.3%; Score 2228; DB 3; Length 2452;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 2312; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
```

```
QY 159 AGGACTATATCCCATTCGCTTTAAATTTAAACAAAGGAGTAATAGGTGAAATCAAC 218
Db 1 AGGACTATATCCCATTCGCTTTAAATTTAAACAAAGGAGTAATAGGTGAAATCAAC 60
QY 219 CCAAGAGCAAGAAAGTTCTTGTTATATTATCGGCGGAGGTGGCTAGAAAGCGCAAG 278
Db 61 CCAAGAGCAAGAAAGTTCTTGTTATATTATCGGCGGAGGTGGCTAGAAAGCGCAAG 120
QY 279 CAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTGCTTACATTAGTGCCCATATTATGG 338
Db 121 CAGAGGGCTTAAAGCTCAACCAACCGGAGCCATTGCTTACATTAGTGCCCATATTATGG 180
```

```
QY 339 ACGAAGCGCGCGTGGAAAAAACCCGTGCCAGCTTATGGAAGAGTGCATGCACTTTT 398
Db 181 ACGAAGCGCGCGTGGAAAAAACCCGTGCCAGCTTATGGAAGAGTGCATGCACTTTT 240
QY 399 TGAAGAAAGATGAAGTAAATGCCCCGGGTGGGTAAATATGTTCCCGATCTAGGTGTAGAAG 458
Db 241 TGAAGAAAGACGAGGTGATGCCCGGGTGGGGAATATGTTCCCTGATTTGGCGTGGGAAG 300
QY 459 CCACCTTTCCTGTATGGTACGAACCTTGTAACTGTGAATTTGGCCCCATCGAACCCAGATGAGC 518
Db 301 CCACCTTTCCTGTATGGTACGAACCTTGTAACTGTGAATTTGGCCCCATCGAACCCAGATGAGC 360
QY 519 ACTTCAAGCGCGGCGAAGTGAATTTGGTTCGGATAAAGACATCGAGCTCAATGCAAGCA 578
Db 361 ACTTAAAGGGGGTGAAGTGAATTTGGCTGTGAATAAGACATTTGAACCTCAACCGCAGTA 420
QY 579 AAGAAGTAAACCGAACTTGAGGTAACTAATGAAGGCGCTAAATCCTTTGCATGTGGGTAGCC 638
Db 421 AGGAAGTTACCGAACTAGAAGTTACTAACGAAGGACCTAAATCCTTTGCATGTGGGTAGCC 480
QY 639 ATTTCCCACTTCTTTGAAGCTAAACAAAGCACTAAATTCGATCGTGAAGAAAGCCTATGGCA 698
Db 481 ATTTCCCACTTCTTTGAAGCCAAACAAAGCATTTGAATTCGATCGGGAAGAAAGCCTATGGCA 540
QY 699 AACGCTTAGATATTCCTCTGCGCAACACGCTACGCAATTTGGGGCAGGACAAACCCGCAAG 758
Db 541 AACGCTTAGATATTCCTCTGCGCAACACCTACGCAATTTGGGGCAGGACAAACCCGCTAAG 600
QY 759 TGCAGTTGATCTCTTTGGTGGCAGTAAAGAGTATTTGGCATGAAACGGGCTTGTGAATA 818
Db 601 TGCAGTTAATCCCTCTTTGGCGGTAGTAAAGAGTATTTGGCATGAAACGGGCTTGTGAATA 660
QY 819 ACATCGCGGATGAACGCATATAACAAAGCGCTTGACAAGCGGCAAAATCTCACGGATTTA 878
Db 661 ATATTGCGAGCAAGCCATATAACAAAGCGCTAGACAAGCAAAATCTCACGGATTTA 720
QY 879 TCAAGTAAAGGAGACTCCCCTGAAATGAAAAACAAGAAATATGTAATACCTTACGAGCCC 938
Db 721 TCAAGTAAAGGAGACTCCCCTGAAATGAAAAACAAGAGTATGTAATACCTTACGAGCCC 780
QY 939 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAATGAC 998
Db 781 ACCAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAATGAC 840
QY 999 TATACCACTATCGCGAAGCACTTAAATTTGGCGCGGTAAACTATCCGTGAGGCTATG 1058
Db 841 TATACCACTATCGCGAAGCACTCAAAATTCGGTGAGGAAACTATCCGTGAGGCTATG 900
QY 1059 GGTCAAGCAATAGCCCTGATGAAAAACACCCCTAGATTTAGTATCATCACTAAACGCGATGATT 1118
Db 901 GGTCAAGCAATAGCCAGATGAAAAACACCTTAGATTTAGTATCATCACTAAACGCGATGATT 960
QY 1119 ATCGACTACACCGGATTTTACAAAGCGGACATTTGGGATTTAAAAACGGCAAAATTCATGGC 1178
Db 961 ATTGACTACACCGGATTTTACAAAGCGGACATTTGGCATTTAAAAATGGAATAATCCATGCG 1020
QY 1179 ATTGGCAAGCGAGAAACAAAGGACATGCAAGATGGCGCTTAAAGCCCTCATATGGTCTGGGT 1238
Db 1021 ATTGGCAAGCGAGAAACAAAGGACATGCAAGATGGCGCTTAAAGCCCTCATATGGTCTGGGT 1080
QY 1239 GTGGGCAAGACGACCTAGCAGGGGAAAGGTATGATTATTACCGCTGGGGGAATCGATTCA 1298
Db 1081 GTGGGCAAGACGACCTAGCAGGGGAAAGGTATGATTATTACCGCTGGGGGAGTCAATTCA 1140
QY 1299 CACACCCACTTCTCTTCTTCCCAACAAATTCCTTACCGCTCTAGCCAAATGGCGTTACAAC 1358
Db 1141 CACACCCACTTCTCTTCTTCCCAACAAATTCCTTACCGCTCTAGCCAAATGGCGTTACAAC 1200
QY 1359 ATGTTTGGAGCGGCAAGGTCTGTAGTAGGCAAGATGCGACTACTATCACTCGGGC 1418
Db 1201 ATGTTTGGCGTGGCAGAGGCCCGGTAGTAGGCAAGATGCGACTACTATCACTCGGGC 1260
QY 1419 AATGGAACCTTGCAACCGCATGTTGCGCGCAGACGAAGAGTATTCTTGAATGTGGGCTTT 1478
```


[illegible]

Db	2341	ATCACCGTCGATCCGAAAACCTTCGAGGCTTTTGTAGATGGCAAACTCTGCACCTCTCTAAA	2400
Qy	2559	CCCACTTCGCAAGTGCCTCTAGCCACGCGCTACACTTTCTTTCTTAGGCACAAT	2610
Db	2401	CCCACTCTGAAGTGCCTCTAGCCCAACGCTACACTTTCTTTCTTAGGCATAAT	2452
<p>RESULT 3</p> <p>US-09-904-994B-4</p> <p>; Sequence 4, Application US/09904994B</p> <p>; Publication No. US20040005325A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: AKZO NO. US20040005325A1el N.V.</p> <p>; TITLE OF INVENTION: Helicobacter vaccine</p> <p>; FILE REFERENCE: Depoedesequities</p> <p>; CURRENT APPLICATION NUMBER: US/09/904,994B</p> <p>; CURRENT FILING DATE: 2001-07-13</p> <p>; NUMBER OF SEQ ID NOS: 21</p> <p>; SOFTWARE: PatentIn Ver. 2.1</p> <p>; SEQ ID NO 4</p> <p>; LENGTH: 2405</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Helicobacter felis</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: (1)..(681)</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: (692)..(2398)</p> <p>US-09-904-994B-4</p>			
<p>Query Match 74.9%; Score 2158.6; DB 3; Length 2405;</p> <p>Best Local Similarity 93.6%; Pred. No. 0;</p> <p>Matches 2251; Conservative 0; Mismatches 154; Indels 0; Gaps 0;</p>			
Qy	206	GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATTATGCGGCGAAGTGGCT	265
Db	1	GTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATATTATGCGGCGAAGTGGCT	60
Qy	266	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT	325
Db	61	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGAAAGCCATTGCTTACATTAGT	120
Qy	326	GCCCATATTATGACGAAGCGCGCGTGGAAAAAACCCTGTCACAGCTTATGGAAGAG	385
Db	121	GCCCATATTATGACGAAGCGCGCGTGGCAAAAACCGTGTCTGNACTTATGGAAGAA	180
Qy	386	TGCATGCATTTTGTAAAAAAGATGAAGTAAATGCCGGGTGGGTAAATATGGTTCCGGAT	445
Db	181	TGTATGCATTTTGTAAAAAAGATGAAGTAAATATGGTGGGAATATGGTCCCTGAT	240
Qy	446	CTAGGTGTAGAGCCACTTTCCTGATGGTACGAAACTGTAACTGTGAATTTGCCCATC	505
Db	241	TTGGGCGGTAGAGCCACTTTCCCGATGGCACCACAACTCGTAAACCGTGAATTTGCCCAT	300
Qy	506	GAAACGATGAGCACTTCAAGCGGGCGAAGTGAAATTTGGTTGCGATAAAGACATCGAG	565
Db	301	GAACTGATGAACAATTTAAAGCGGTGAAGTGAATTTGGCTGTGATAAAGACATTGAG	360
Qy	566	CTCAATGCAGGCAAAAGAAAGTAAACCGAACTTTGAGGTACTAATGAAGGCGCTAAATCCTTG	625
Db	361	CTCAACGCGGTAAGGAAGTTACCGAGCTTTGAAAGTTACCAACGAAGGACCTAAATCCTTG	420
Qy	626	CATGTGGTAGCCATTTCCACTTCTTTGAAGCTAAACGAAGCACTAAATTCGATCGTGA	685
Db	421	CATGTGGTAGCCATTTCCACTTCTTTGAAACCAACGAAGGCACTTGAATTTGATTCGGGAA	480
Qy	686	AAAGCCTATGCAAAACGCTAGATATTCCCTCTGGCAACACGCTTACGCATTTGGGCGAGGA	745
Db	481	AAAGCCTATGCAAAACGCTAGATATTCCCTCTGGCAACACGCTTACGCATTTGGGCGAGGA	540
Qy	746	CAAAACCGCGAAAGTGCAGTTGATTCCCTCTTTGGTGGCAGTAAAAAGTGAATTGGCATGAAC	805

Db 541 CAAACCCGTAAGTGCAGTTAATCCCTCTTTGGCGGTAGTAAAGATGATGGCATGAAC 600
Qy 806 GGGCTTTGTGAATAACATCGCGATGAACGCATATAAATGAGCTGTGCAAGGGGAAA 865
Db 601 GGGCTTTGTGAATAATTTGGGACGAAAGCCATATAACACAAAGCACTAGACAAAGGCAAAA 660
Qy 866 TCTACGGATTTATCAAGTAAGAGAGACTCCCATGAAATAAGAAAAACAAGAAATATGTAA 925
Db 661 TCTACGGATTTATCAAGTAAGAGAGACTCCCATGAAATAAGAAAAACAAGAGTATGTAA 720
Qy 926 TACCTACGACCCACCAAGGGGATAAAGTCCCTTAGGAGATACCGATCTTTGGGCGA 985
Db 721 CACCTACGACCCACCAAGGGGATAAAGTCCCTTAGGAGATACCGATCTTTGGGCGA 780
Qy 986 AGTAGAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAACATAT 1045
Db 781 AGTAGAATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGCGGTAAACATAT 840
Qy 1046 CCGTAGGGGTATGGGTAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATCAC 1105
Db 841 CCGTAGGGGTATGGGTAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCATCAC 900
Qy 1106 TAAACGGATGATATGATACACCGGATTTTAAAGCCGACATTTGGGATTTAAACGG 1165
Db 901 CAAACGGATGATATGATACACCGGATTTTAAAGCCGACATTTGGGATTTAAACGG 960
Qy 1166 CAAATCCATGGCATTTGGAAGCGGGAACAAAGGACATGCAAGATGGGTAAGCCCTCA 1225
Db 961 CAAATCCATGGCATTTGGAAGCGGGAACAAAGGACATGCAAGATGGGTAAGCCCTCA 1020
Qy 1226 TATGGTCTGGGTGTGGGACAGAACCTAGCAGGGGAGGTATGATTTATTCACGGTGG 1285
Db 1021 TATGGTCTGGGTGTGGGACAGAACCTAGCAGGGGAGGTATGATTTATTCACGGTGG 1080
Qy 1286 GGAATCGATTCACACCCCATCTTCTTCTTCCACAAATTCCTACCGCTTAGCCAA 1345
Db 1081 GGGATCGATTCACACCCCATCTTCTTCTTCCACAAATTCCTACCGCTTAGCCAA 1140
Qy 1346 TGGCGTTACAAACATGTTTGGAGCGGCACAGTCTCTGATGAGGACCAAGTCCGACTAC 1405
Db 1141 TGGTGTACAAACATGTTTGGAGTGGCACAGGTCCGCTGATGAGGACCAAGTCCGACTAC 1200
Qy 1406 TATCACTCCGGGCAATGGAATTTGACACCGCATGTTGCGGCGAGCAGAGAGATTTCTAT 1465
Db 1201 CATCACTCCGGGCAATGGAATTTGACACCGCATGTTGCGGCGAGCTGAGAGATTTCTAT 1260
Qy 1466 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1525
Db 1261 GAATGTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGAACAGT 1320
Qy 1526 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCCGAT 1585
Db 1321 AGAAGCGGGCGGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAAACCAAGTCCGAT 1380
Qy 1586 CGATCACTCTTGAGCGTGCAGATCAATACGATGTCAGATTTGTATCCACCGCATAC 1645
Db 1381 CGATCACTCTTGAGCGTGCAGATCAATACGATGTCAGATTTGTATCCACCGCATAC 1440
Qy 1646 AGTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGCAATGAAAGCGGCGCCATCCA 1705
Db 1441 GGTCAATGAGGCGAGTTATGTAGATGACACCCCTAAATGCAATGAAAGCGGCGCCATCCA 1500
Qy 1706 TGCCTACCAATTAGGAGGCGGGTGGAGGACATCACTCATGTATGTTATCAACATGGCAGG 1765
Db 1501 TGCCTACCAATTAGGAGGCGGGTGGAGGACATCACTCATGTATGTTATCAACATGGCAGG 1560
Qy 1766 CGAGCTCAATATTTCTACCTCTCCGACCCCACTATTTCCCTATACCATTAATACGGT 1825
Db 1561 CGAGCTCAATATTTCTACCTCTCCGACCCCACTATTTCCCTATACCATTAATACGGT 1620
Qy 1826 TGCAGAACACTTAGACATGCTCATGATGCTGCGCACCCCTAGACAAAGCAATCCGCGAGG 1885
Db 1621 TGCAGAACACTTAGACATGCTCATGATGCTGCGCACCCCTAGACAAAGCAATCCGCGAGG 1680

Qy 1886 TTTACAAATTTTCTCAAAAGCCGATATCCGCCCCGGCTCTATTCGGGCTGAAGATGTGCTCCA 1945
Db 1681 TCTCCAGTTTTCCTCAAAAGCCGATATCCGCCCCGGCTCTATTCGGCTGAAGATGTGCTCCA 1740
Qy 1946 TGATATGGGTGTGATCGCGATGACAAAGCTCCGATTTCCGAAGCAATGGGCGCTGCGAGCA 2005
Db 1741 TGATATGGGTGTGATCGCGATGACAAAGCTCCGATTTCCGAAGCAATGGGCGCTGCGAG 1800
Qy 2006 AGTGATTTCTCGAACTTTGGGACAGCTCGCGATTAAGAAATAAAAAAGAAATTTGGTAAGCTTCC 2065
Db 1801 AGTGATTTCTAGAACTTTGGGCAAACTGACAGACAGAAATAAAAAAGAAATTTGGTAAGCTTCC 1860
Qy 2066 TGAAGATGGCAAGATAAACGATAATTTCCGCAATTAAGCGCTACATCTCCTCAAAATACACTAT 2125
Db 1861 TGAAGATGGTGCAGATAATGACAACTTCGCGATCAAAAGCTATATCTCCTCAAAATACACCAT 1920
Qy 2126 CAACCCCGCTTTGACCCACCGCGTGAAGAGTATATCGGCTCTGTTGGAAGAGGCAAGAT 2185
Db 1921 TAAATCCCGCTTTGACCCATGCGCGTGAAGAGTATATCGGCTCTGTTGGAAGAGGCAAGAT 1980
Qy 2186 CGCGACTTGGTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCAAAATCGTGATCAA 2245
Db 1981 CGCGACTTGGTGTGGAAATCTGCTTTTTTTGGCGTAAAAACCAAAATCGTGATCAA 2040
Qy 2246 AGCGGTATGGTGTCTTCTCTGAAAATGGGCGATTTCTAACCGCTCTGTGCCCATCTCCCA 2305
Db 2041 AGCGGTATGGTGTCTTCTCTGAAAATGGGCGATTTCTAACCGCTCTGTGCCCATCTCCCA 2100
Qy 2306 ACCGTTTTATTAACCGGAAATTTTGGGCAATCAAGGCAAGGCAAAATTTGACACCAAGAT 2365
Db 2101 GCGGTTTTATTAACCGGAAATTTTGGGCAATCAAGGCAAGGCAAAATTTGACACCAAGAT 2160
Qy 2366 CACTTTTTTTTCCAAAGCTCGCTATGAAAATGGGCGTGAAGAAAGCTGGGCTTAGAGCG 2425
Db 2161 CACTTTTTTTTCCAAAGCTCGCTATGAAAATGGGCGTGAAGAAAGCTGGGCTTAGAGCG 2220
Qy 2426 CCAAGTTCTACCGGTCAAAATCTGCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGA 2485
Db 2221 CAAGGTCTACCGGTGAAGAAATCTGCGGCAACATCACTAAGAAAGACTTCAAAATCAACAA 2280
Qy 2486 CAAACCGGCAAAATCAACCGTCTGATCGGAAACCTTCGAGGTCTTTTGTAGATGGCAAACT 2545
Db 2281 CAAGCGGCGCATATCACTGTGATCTTAAACCTTAAACCTTTCGAGGTCTTTGTAGATGGCAAACT 2340
Qy 2546 CTGCACTCTTAAACCGCACTCGCAAGTCCCTTAGCCCGAGCGCTACATTTTCTCTAGCG 2605
Db 2341 CTGCACTCTTAAACCGCGCTCTGAAGTCTCTAGCCCAACGCTACACTTTTCTCTAGCG 2400
Qy 2606 ACAAT 2610
Db 2401 ACAAT 2405

RESULT 4
US-09-904-994B-10
; Sequence 10, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(682)

; FEATURE:									
; NAME/KEY: CDS									
; LOCATION: (693) .. (2399)									
US-09-904-994B-10									
Query Match 74.9%; Score 2158.6; DB 3; Length 2407;									
Best Local Similarity 93.6%; Pred. No. 0;									
Matches 2251; Conservative 0; Mismatches 155; Indels 0; Gaps 0;									
Qy	206	GTGAAGTCTACACCCCAAGAGAGGAGTTCCTGTATATATATGCGGGCGAAGTGGCT	265	1166	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAGACATGCAAGATGCGCTAAGCCCTCA	1225	Qy	1166	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAGACATGCAAGATGCGCTAAGCCCTCA
Db	2	GTGAAGTCTACACCCCAAGAGAGGAGTTCCTGTATATATATGCGGGCGAAGTGGCT	61	962	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAGACATGCAAGATGCGCTAAGCCCTCA	1021	Db	962	CAAAATCCATGGCATTTGGCAAGGCAAGAAACAAGACATGCAAGATGCGCTAAGCCCTCA
Qy	266	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT	325	1226	TATGGTCTGGGTGTGGGCACAGACACTAGCAGGGGAAGGTATGATTATTCACCGCTGG	1285	Qy	1226	TATGGTCTGGGTGTGGGCACAGACACTAGCAGGGGAAGGTATGATTATTCACCGCTGG
Db	62	AGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAAGCCATTGCTTACATTAGT	121	1022	TATGGTCTGGGTGTGGGCACAGACACTAGCAGGGGAAGGTATGATTATTCACCGCTGG	1081	Db	1022	TATGGTCTGGGTGTGGGCACAGACACTAGCAGGGGAAGGTATGATTATTCACCGCTGG
Qy	326	GCCCATATATGACGAAAGCGGCCGTGGAAAAAACCCTGGTCCAGCTTATGGAAGAG	395	1286	GGGAATCGATTTCACACCCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA	1345	Qy	1286	GGGAATCGATTTCACACCCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA
Db	122	GCCCATATATGACGAAAGCGGCCGTGGAAAAAACCCTGGTCCAGCTTATGGAAGAG	181	1082	GGGGATCGATTTCACACCCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA	1141	Db	1082	GGGGATCGATTTCACACCCCACTTCTTCTCCACAACTTCCCTACCGCTCTAGCCAA
Qy	386	TCATGCACTTTTGAAGAAAGATGAAGTAATGSCCGGGTGGTAAATGTTCCCGAT	445	1346	TGCGGTTTAAACCATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGCGACTAC	1405	Qy	1346	TGCGGTTTAAACCATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATGCGACTAC
Db	182	TGATGCACTTTTGAAGAAAGATGAAGTAATGSCCGGGTGGTAAATGTTCCCGAT	241	1142	TGCGGTTTAAACCATGTTTGGCGTGGCACAGGTCGCTAGATGCGCAAGATGCGACTAC	1201	Db	1142	TGCGGTTTAAACCATGTTTGGCGTGGCACAGGTCGCTAGATGCGCAAGATGCGACTAC
Qy	446	CTAGGTGTAGAAGCCACCTTTCTGATGTAAGAACTTGTAACTGTGAATTGGCCCATC	505	1406	TATCACTCCGGGCAATTGGAACTTGGACCGCATGTTGCGGCAGCAGAGAGTATTCTAT	1465	Qy	1406	TATCACTCCGGGCAATTGGAACTTGGACCGCATGTTGCGGCAGCAGAGAGTATTCTAT
Db	242	TTAGGGGTGGAAGCTACTTTTCCGATGGCACCAACTCGTAACCGTGAATTGGCCCATC	301	1202	CATCACTCCGGGCAATTGGAACTTGGACCGCATGTTGCGGCAGCTGAAGAGTATTCTAT	1261	Db	1202	CATCACTCCGGGCAATTGGAACTTGGACCGCATGTTGCGGCAGCTGAAGAGTATTCTAT
Qy	506	GAAACGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAG	565	1466	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT	1525	Qy	1466	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT
Db	302	GAAACGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAAGACATCGAA	361	1262	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT	1321	Db	1262	GAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGTAGAACAACT
Qy	566	CTCAATGCAAGCAAGCACTTGAAGTACCGAATGAGGTTACTAATCAAGGCTTAATCTTG	625	1526	AGAAGCGGGCGGATTTGGTTTAAATTTGATGAAGACTTGGGGCACACACCAAGTGCAT	1585	Qy	1526	AGAAGCGGGCGGATTTGGTTTAAATTTGATGAAGACTTGGGGCACACACCAAGTGCAT
Db	362	CTCAATGCAAGCAAGCACTTGAAGTACCGAATGAGGTTACTAATCAAGGCTTAATCTTG	421	1322	AGAAGCGGGCGGATTTGGTTTAAATTTGATGAAGACTTGGGGCACACCTCAAGTGAAT	1381	Db	1322	AGAAGCGGGCGGATTTGGTTTAAATTTGATGAAGACTTGGGGCACACCTCAAGTGAAT
Qy	626	CATGTGGGTAGGCAATTTCCACTTCTTTGAAGCTTAACAAGGCACATAAAATTCGATCGTGA	685	1586	CGATCACTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1645	Qy	1586	CGATCACTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC
Db	422	CATGTGGGTAGGCAATTTCCACTTCTTTGAAGCCACACAGGCATTGAATTCGATCGGAA	481	1382	CGATCACTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC	1441	Db	1382	CGATCACTGTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATAC
Qy	686	AAAGCCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTACGCAATGGGGCAGGA	745	1646	AGTCAATGAGCGAGTTTATGTAGATGACACCTCAATGCAATGAACGGCGCGCCATCCA	1705	Qy	1646	AGTCAATGAGCGAGTTTATGTAGATGACACCTCAATGCAATGAACGGCGCGCCATCCA
Db	482	AAAGCCTATGGCAACCGCTAGATATTCCTCTGGCAACACGCTACGCAATGGGGCAGGA	541	1442	GGTCAATGAGCGAGTTTATGTAGATGACACCTCAATGCAATGAACGGCGCGCCATCCA	1501	Db	1442	GGTCAATGAGCGAGTTTATGTAGATGACACCTCAATGCAATGAACGGCGCGCCATCCA
Qy	746	CAAAACCGCAAGTGCATGATTCTCTTGGTGGCAGTAAAGAGTATTGCGCATGAAC	805	1706	TGCTTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTATTACCATGGCAGG	1765	Qy	1706	TGCTTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTATTACCATGGCAGG
Db	542	CAAAACCGTAAAGTGCATGATTCTCTTGGCGGCAAGTAAAGAGTATTGCGCATGAAC	601	1502	TGCTTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTATTACCATGGCAGG	1561	Db	1502	TGCTTACCAATTTGAGGGAGCGGTGGAGGACACTCACCTGATGTATTACCATGGCAGG
Qy	806	GGGCTTGTGAATAACATCGCGATGAACGCATATAACATATAAGCGCTTGACAGGCGGAA	865	1766	CGAGCTCAATATTTCTACCTCTCCACACCCCCCACTATTTCCTTATACCATTAATACGGT	1825	Qy	1766	CGAGCTCAATATTTCTACCTCTCCACACCCCCCACTATTTCCTTATACCATTAATACGGT
Db	602	GGGCTTGTGAATAATTTGAGATGAACGCATATAACATATAAGCGCTTGACAGGCGGAA	661	1562	CGAGCTCAATATTTCTACCTCTCCACACCCCCCACTATTTCCTTATACCATTAATACGGT	1621	Db	1562	CGAGCTCAATATTTCTACCTCTCCACACCCCCCACTATTTCCTTATACCATTAATACGGT
Qy	866	TCTACGGATTTATCAAGTAAGAGACTCCCATGAAATTAAGAAAAACAAGATATGTAA	925	1826	TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAAACGATCCGCGAGA	1885	Qy	1826	TGCAGAACACTTAGACATGCTCATGACATGCCACCACTAGACAAAACGATCCGCGAGA
Db	662	TCTACGGATTTATCAATTAAGAGACTCCCATGAAATTAAGAAAAACAAGATATGTAA	721	1622	TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTAGATAAAGCATCCCGGAGA	1681	Db	1622	TGCAGAACACTTAGACATGCTCATGACCTGCCACCACTAGATAAAGCATCCCGGAGA
Qy	926	TACCTACGACCCCAACAAAGGCGATTAAGTGCCTTAGGAGATACCGATCTTTTGGGCGA	985	1886	TTTACAAATTTTCTCAAGCCGTTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA	1945	Qy	1886	TTTACAAATTTTCTCAAGCCGTTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA
Db	722	TACCTACGACCCCAACCAAGGCGATTAAGTGCCTTAGGAGATACCGATCTTTTGGGCGA	781	1682	TTTACAAATTTTCTCAAGCCGTTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA	1741	Db	1682	TTTACAAATTTTCTCAAGCCGTTATCCCGCCCGGCTCTATCGCGCTGAAGATGTGCTCCA
Qy	986	AGTGAACATGACTATACCACTTATCGGAGAGCACTTAAATTTGGCGGGTAAACTAT	1045	1946	TGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCA	2005	Qy	1946	TGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCA
Db	782	AGTGAACATGACTATACCACTTATCGGAGAGCTTAAATTTGGCGGGTAAACTAT	841	1742	TGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCA	1801	Db	1742	TGATATGGGTGTGATCCGATGACAAAGCTCGGATTCGCAAGCAATGGGGCGTGCAGGCA
Qy	1046	CGGTAGGGTATGGGTGAGAGCAATAGCCCTGATGAAACACCCCTAGATTAGTCAATCAC	1105	2006	AGTGAATTCCTCGAACTTGGCAGACTGCGGATAGAAATAAAGAAATTTGGTAAGCTTCC	2065	Qy	2006	AGTGAATTCCTCGAACTTGGCAGACTGCGGATAGAAATAAAGAAATTTGGTAAGCTTCC
Db	842	CGGTAGGGTATGGGTGAGAGCAATAGTCCAGATGAAACACCCCTAGATTAGTCAATCAC	901	1802	AGTGAATTCCTCGAACTTGGCAGACTGCGGATAGAAATAAAGAAATTTGGTAAGCTTCC	1861	Db	1802	AGTGAATTCCTCGAACTTGGCAGACTGCGGATAGAAATAAAGAAATTTGGTAAGCTTCC
Qy	1106	TAAACCGATGATTATCGATACACCGGATTTTCAAGCGGACATTTGGGATTTAAAAACGG	1165	2066	TGAAGATGGCAAGATACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT	2125	Qy	2066	TGAAGATGGCAAGATACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT
Db	902	CAACGGGATGATTATTGACTTACACCGGATTTTCAAAAGCGGACATTTGGGATTTAAAAATGG	961	1862	TGAAGATGGCAAGATACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT	2121	Db	1862	TGAAGATGGCAAGATACGATTAATTTCCGATTAAGCGCTACATCTCCAATACACTAT

```
Db      2042 AGCGGTATGGTGGTCTTCTCTGAAATGGCGACTCAACCGCGTCGCTACACCTCA 2101
Qy      2306 ACCGGTTTATTACCGGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACCGCAT 2365
Db      2102 GCCGGTTTATTACCGGAAATGTTTGGGCATCAGCGCAAGCGGAAATTTTGACACCGCAT 2161
Qy      2366 CACTTTTGTGTTCCAAAGTCGCCTATGAAAAATGGCGTGAAAGAAAAGCTGGCGTTAGAGCG 2425
Db      2162 CACTTTTGTGTTCCAAAGTCGCCTATGAAAAATGGCGTGAAAGAAAAGCTAGGCTTAGAGCG 2221
Qy      2426 CCAAGTTTACCGGTCAAAATCGCGTAACATCAACAAGAAAGACTTCAAGTTCAACGA 2485
Db      2222 CAAAGTGCTACCGGTGAAAAACTGCCGCAACATCACTAAGAAAGACTTCAAAATTCACAA 2281
Qy      2486 CAAAACGGCAAAATCACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGCAAACT 2545
Db      2282 CAAAGCGGCGCATATCACTGTCATCTTAAACCTTCGAGGTCTTTGTAGATGCAAACT 2341
Qy      2546 CTGCACTCTTAAACCCACCTCGCAAGTGCTCTAGCCAGCGCTACACTTTTCTTAGGC 2605
Db      2342 CTGCACTCTTAAACCCGCTCTGAAAGTGCTCTAGCCAGCGCTACACTTTTCTTAGGC 2401
Qy      2606 ACAATG 2611
Db      2402 NCAATG 2407

RESULT 5
US-09-904-994B-7
; Sequence 7, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2183
; TYPE: DNA
; ORGANISM: Helicobacter felis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(683)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (694)..(2181)
; US-09-904-994B-7

Query Match      67.6%; Score 1948; DB 3; Length 2183;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 2035; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy      206 GTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGCT 265
Db      3 GTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGTATATATATGCGGCGAAGTGCT 62
Qy      266 AGAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGCAAGCCATTGCTTACATTAGT 325
Db      63 AGAAGCGCAAGCAGAGGCTTAAAGCTCAATCAACCCGCAAGCCATTGCTTACATTAGT 122
Qy      326 GCCCATATTATGACCAAGCGCGCTGGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385
Db      123 GCCCATATTATGACCAAGCGCGCTGGGAAAAAACCCTTGCTGAACTTATGGAAGAA 182
Qy      386 TGCATGCACCTTTTGAAGAAAGATGAAGTAATCCCGGGGTGGGTAAATATGGTTCCCGAT 445
Db      183 TGTATGCACCTTTTGAAGAAAGATGAGGTGATGCCCGGTGGGGAATATGGTCCCTGAT 242
Qy      446 CTAGGTGTAGAACCCACTTTCCTGATGGTACGAACTTGTAACTGTGAATTGGCCCATC 505
```

```
Db      243 TTGGCGGTAGAACCACTTTCCCGATGGCACCAAACTCGTAACCGTGAATTTGGGCCATT 302
Qy      506 GAAACAGATGAGCACTTCAAAAGCGGGCGAAGTGAATTTGGTTGGATAAAGACATCGAG 565
Db      303 GAACCTGATGACACTTTAAAGCCGCTGAAGTGAATTTGGCTGTGATAAAGACATTGAG 362
Qy      566 CTCAATGAGGCAAGAGTAACCGAACTTGAGGTTACTTAATGAAGGGGCTAAATCCTTG 625
Db      363 CTCAACGTTGGGTAAGGAAGTTACCGAGCTTGAAGTTACCAACGAAGGACCTAAATCCTTG 422
Qy      626 CATGTGGGTAGCACTTTCCACTTCTTTGAAGCTAAACAGGCACTAAATTCGATCGTGA 685
Db      423 CATGTGGGTAGCACTTTCCACTTCTTTGAAACCAACAAAGGCACTGAAATTCGATCGGAA 482
Qy      686 AAAGCCTATGCAAAACGGCTAGATATTCCTCTCTGGCAACACGCTACGCACTGGGGCAGGA 745
Db      483 AAAGCCTATGCAAAACGGCTAGATATTCCTCTCTGGCAACACGCTACGCACTGGGGCAGGA 542
Qy      746 CAAACCCGCAAGTGCAGTTGATTCCTCTCTGGTGGCAGTAAAAAAAGTGAATTGGCAATGAAC 805
Db      543 CAAACCCGTAAGTGCAGTTAATCCCTCTTGGCGGTAGTAAAAAAAGTGAATTGGCAATGAAC 602
Qy      806 GGGCTTGTGAATAACATCGCGGATGAACCGCATAAACATAAAGCGCTTGACAAAGGCGAAA 865
Db      603 GGGCTTGTGAATAATAATTGCGGACGAACGCCATAAACAACAAAGCACTAGACAAAGGCAAAA 662
Qy      866 TCTCACCGATTTATCAAGTAGGAGACTCCCATGAATGAATAAACAACAGAAATATGTAAA 925
Db      663 TCTCACCGATTTATCAAGTAGGAGACTCCCATGAATGAATAAACAACAGAAATATGTAAA 722
Qy      926 TACCTACGGAACCCACCAAGCGCATAAAGTGCCTTAGGAGATACCGCATCTTTTGGGCAGA 985
Db      723 CACCTACGGAACCCACCAAGCGCATAAAGTGCCTTAGGAGATACCGCATCTTTTGGGCAGA 782
Qy      986 AGTAGAACATGACTATATACCCTATGCGGAAGAACTTAAATTTGGCGGGTAAAACTAT 1045
Db      783 AGTAGAACATGACTATATACCCTATGCGGAAGAGCTCAAAATTTGGCGGGTAAAACTAT 842
Qy      1046 CCGTAGGAGGTATGGTCAGAGCAATAGCCCTGATGAATAAACAACCCCTAGATTAGTCATCAC 1105
Db      843 CCGTAGGAGGTATGGTCAGAGCAATAGCCCAATAGCCCAATAGATAATTTAGTATGATCAC 902
Qy      1106 TAACGCGATGATTATCGACTACACCGGATTTTACAAAGCCGATTTGGCATTTAAAAACGG 1165
Db      903 CAACGCGATGATTATCGACTACACCGGATTTTAAAGCCGACATTTGGTTTAAAAATGG 962
Qy      1166 CAAAATCCATGGCATTTGCAAGGCGAGGAAACAAAGGACATGCAAGATGGCGTAAGCCCTCA 1225
Db      963 CAAAATCCATGGTATTGGCAAGCGGGGAAACAAAGACATGCAAGATGGCGTAAGCCCTCA 1022
Qy      1226 TATGCTGTGGGTGTGGGCACAGAAAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1285
Db      1023 TATGCTGTGGGTGTGGGCACAGAAAGCACTAGCAGGGAAGGTATGATTTATACCGCTGG 1082
Qy      1286 GGGATCGATTACACACCCACTTCTTCTCCCAAACTTCCCTACCGCTCTAGCCAA 1345
Db      1083 GGGATCGATTTCGACACCCACTTCTCTCCCAAACTTCCCTACCGCTCTAGCCAA 1142
Qy      1346 TGGGTTTCAACCACTGTTTGGAGCGGCACAGGTCTCTAGATGGCAGCAATTCGCACTAC 1405
Db      1143 TGGTGTACACCACTGTTTGGAGGTGGCAGAGTCCGGTAGATGGCAGCAATTCGCACTAC 1202
Qy      1406 TATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTGCGCGCAGCAGAAAGATTTCTAT 1465
Db      1203 CATCACTCCGGGCAAAATGGAACTTGCAACCGCATGTTGCGCGCAGCTGAAGAGTATTCTAT 1262
Qy      1466 GAATGTGGCTTTTGGGCAAAAGCAATAGCTCTAGCAAAACCACTTGTAGAACAACT 1525
Db      1263 GAATGTGGCTTTTGGGCAAAAGCAATAGTCTTAGCAAAACCACTTGTAGAACAACT 1322
Qy      1526 AGAAGCGGCGCATTTGGTTTTAAATTCGATGAAGACTGGGGCAACCAACCAAGTCGAT 1585
```

Db 1323 AGAAGCGGCGGATGGCTTTAAATTGATGAAGACTGGGCGACAAACCAAGTCCGAT 1382
QY 1586 CGATCACTGCTTTAGCGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATAC 1645
Db 1383 CGATCACTGCTTTAGCGTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATAC 1442
QY 1646 AGTCAATGAGGAGGTTATGTAGATGACACCTAAATGAATGAACGGGCGGCGCATCCA 1705
Db 1443 GGTCAATGAGGAGGTTATGTAGATGACACCTAAATGAATGAACGGGCGGCGCATCCA 1502
QY 1706 TGCCTTACCACATTTAGGAGGCGGTTGGAGGACACTCAGCTGATGTTTATCACCATGCGAGG 1765
Db 1503 TGCCTTACCACATTTAGGAGGCGGTTGGAGGACACTCAGCTGATGTTTATCACCATGCGAGG 1562
QY 1766 CGAGCTCAATATTTCTCAAGCCGTTATCCACACCCCGCTATTTCCCTATACCATTAATACGGT 1825
Db 1563 CGAGCTCAATATTTCTCAAGCCGTTATCCACACCCCGCTATTTCCCTATACCATTAATACGGT 1622
QY 1826 TGCAGAACACTTAGACATGCTATGACATGCGACACCTAGACAAACGGATCCGCGAGGA 1885
Db 1623 TGCAGAACACTTAGACATGCTATGACATGCGACACCTAGACAAACGGATCCGCGAGGA 1682
QY 1886 TTTTACAATTTTCTCAAGCCGTTATCCGCGCTCTATCGCGCTGAGATGTCCTCA 1945
Db 1683 TTTTACAATTTTCTCAAGCCGTTATCCGCGCTCTATCGCGCTGAGATGTCCTCA 1742
QY 1946 TGATATGGGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGCTGACGCGA 2005
Db 1743 TGATATGGGTGATCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGCTGACGCGA 1802
QY 2006 AGTGATTCCTGAACTTGGCAGACTCGCGATAGGAATAAAGAAATTTGGTAAAGCTTC 2065
Db 1803 AGTGATTCCTGAACTTGGCAGACTCGAGATAGGAATAAAGAAATTTGGTAAAGCTTC 1862
QY 2066 TGAAGATGCAAGATGAAGTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 2125
Db 1863 TGAAGATGCAAGATGAAGTAATTTCCGATTAAGCGCTACATCTCCAAATACACTAT 1922
QY 2126 CAACCCGCTTTGACCCACCGGCTGAGGATATATCGGCTCTGTGGAAGAGGGCAAGAT 2185
Db 1923 TAATCCCGCTTTGACCCACCGGCTGAGGATATATCGGCTCTGTGGAAGAGGGCAAGAT 1982
QY 2186 CGCGACTTGGTGGTGGAACTCTGCTTTTGGCGTAAACCCAAATTCGTGATCAA 2245
Db 1983 CGCGACTTGGTGGTGGAACTCTGCTTTTGGCGTAAACCCAAATTCGTGATCAA 2042
QY 2246 AGCGGTATGGTGGTCTTCTGAAATGGGCGATTTCTAACGCTCTGTGCGCCACTCCCCA 2305
Db 2043 AGGTGGCATGGTGGTCTTCTGAAATGGGCGATTTCTAACGCTCTGTGCGCCACTCCCCA 2102
QY 2306 ACCGGTTTATACCGGAAATGTTGGGATCAGCGGCAAGGCGAAATTTGACACCGCAT 2365
Db 2103 GCGGTTTATACCGGAAATGTTGGGATCAGCGGCAAGGCGAAATTTGACACCGCAT 2162
QY 2366 CACTTTTGTTCCTCAAGTCG 2385
Db 2163 CACTTTTGTGTCCTCAAGCG 2182

RESULT 6

US-10-639-273-1
; Sequence 1, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639, 273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337

; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8407
; TYPE: DNA
; ORGANISM: Helicobacter bizzoeronii
US-10-639-273-1

Query Match 41.1%; Score 1184.4; DB 7; Length 8407;
Best Local Similarity 68.5%; Pred. No. 7.3e-260;
Matches 1681; Conservative 0; Mismatches 741; Indels 32; Gaps 2;

QY 183 AAATTAAACAAGGAGTAAAGTGAAACTCACACCCAAAGAGCAAGAAAGTCTTGT 242
Db 2319 AAATTGGTAGAAGGAGTTAGGATGAAATTAACCCCTAAAGAGCTGACAAAGTCTGT 2378
QY 243 TATATATGCGGCGAAGTGGCTAGAAAGCAAGCAGAGGGCTTAAAGCTCAACCAAC 302
Db 2379 TGCAATTATGCGGCGAATTTGGCTAATAAAGCAAGCAAAATGGCGTTAAAGCTAAATATA 2438
QY 303 CCGAAGCCATTGCTTACATTAGTCCCATATTTATGGAAGAGCGCGCTGGAAAAAAA 362
Db 2439 CTGAGGAGTAGCCCTCATCAGTCCCTGTGTGGAAGAGCCCGTGCAGGTAAAAAAA 2498
QY 363 CCCTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGAAGAAAGATGAAGTAATGCCCG 422
Db 2499 GTGTGGCGGATTTGATGCAAGAGGAGGACAGCACATTTCTAAAGCTGATGATGTCATGCCCG 2558
QY 423 GGTGGGTAAATATGTTTCCCGATCTAGGTGTAGAGCCACCTTTCTGTAGTGGTAGAAAC 482
Db 2559 GTGTAGCCCATATGATCCACGAAGTGGGATTTGAAGCTAACTTCCCTGATGGGACAAAAC 2618
QY 483 TTGTAAGTGAATTTGGCCCATCGAACAGATGAGCACTTCAAGCGGCGGAGTGAAT 542
Db 2619 TGTAACCATCCATACCCCGTTGAAGATGGTGGCATAAATGGCTCCGGTGAAGTGA 2678
QY 543 TTGGTTGCGATAAAGACATCGAGCTCAATCAGGCAAAAGATGAACCGAACTTGAGGTGA 602
Db 2679 TTTTGAAGAACGAAGACATCACTTTGAATGAGCGCAAAAGCCACCACTTTAGAGTGC 2738
QY 603 CTAAATGAAGGCGCTAAATCTTGTGATGTTGGTAGCCATTTTCCACTTTCTTGAAGTAAACA 662
Db 2739 ATAAACAAAGCGATCGCCCGTGCAGTGGGCTCCCACTTCCACTTTCTTGAAGTGAATA 2798
QY 663 AGCAGCTAAATTCGATCGTGAAGAAACCTATGCGCAACCGCTAGATATTTCCCTCTGCGA 722
Db 2799 AGCTTTTGAATTTGATTCGTGAAGAAAGCCTATGCGCAACCGCTAGACATTTCTCTGGAA 2858
QY 723 ACAGCTACGCAATTTGGGCGAGGACAAACCCGAAAGTGCAGTTGATTTCTCTTGGTGGCA 782
Db 2859 CCCTGTGCGCTTTGAACCGGTGAGAAAAAACCCGTGAATTTGATTTCAATTTGGCGTGA 2918
QY 783 GTAAAAAGTGTATGGCATGAACGGGCTTTGTGAATAACATCGCGATGAACCGCAATAAAC 842
Db 2919 ACCAAGCATTTAGCGGCTTTAACTCTCTTGTGGATCGCAAGCGGATCTGATGSCAAA 2978
QY 843 ATAAAGGCTTTGACAGGCGAAATCTCAGGATTTT----- 877
Db 2979 AACTTGTCTCTCAACCGCGCAAGAAACATGGCTTTTGGTGTGTGAATTTGCGGTTGCGGATA 3038
QY 878 -----ATCAAGTAAAGGAGCTCCCATGAAATGAAAAA---ACAAGATATCTAAATACCT 930
Db 3039 AAAAAATAGGAAAGGACAAATCCGATGAAAAAAATCTCTCGAAAAAGAAATATGTTTCTATGT 3098
QY 931 ACGGACCCACCAAGGCGATAAAGTGGCTTTAGGAGATACCGATCTTTGGGCGAGAGTAG 990
Db 3099 ATGAGCCCACTACGGGCGATAAAGTGAATTTGGCGGATACCGACTGATCTTAGAAGTCG 3158
QY 991 AACATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGCTTAAACTATTCCTGTG 1050
Db 3159 AACATGACTGACCCACTTATGGCGGAGAAATTAAGTTTGGTGGCGGTAAACCACTTCGCG 3218

Db 187 AGCCAGAACTCGATCTTGTGCTCACTAAGCCCTGATCGTGGATTATACACCGGCATTTAT 246
Qy 1140 AAAGCCGATTTGGGATTTAAAAACGCAAAATCCATGGCATTTGGCAAGCAGCAAGAAACAG 1199
Db 247 AAAGCCGATTTGGGATTTAAAAATGGCAAAATCCATGGCATTTGGCAAGCAGCAAGAAACAG 306
Qy 1200 GACATGCAAGATGGCGTGAAGCCCTCATATGGTCTGGTGGTGGGACAGCAAGCACTAGCA 1259
Db 307 GACATGCAAGATGGCGTGGTGAAGCAATCTTTGGTGGGCCCTGCTACTGAGGCTTTGGCC 366
Qy 1260 GGGGAAGGTATGATTTATACCCCTGGGGAATCGATTACACACCCACATTTCTTTCTCCA 1319
Db 367 GCTGAAGGCTGATTTGATACAGCTGGTGGATTGACACCCATTCACATCTTTATTTCTCCC 426
Qy 1320 CAACAATTTCCCTACCGCTCTAGCCCAATGGCGTTTAAACCATGTTTGGAGCGGCACAGGT 1379
Db 427 CAACAATTTCCACACAGCATTTGCCAGCGGGATCAACAACATGATTTGGTGGGGACAGGT 486
Qy 1380 CTTGTAGATGGCAAGATGCGACTACTATCATCTCGGGCAAAATGGAACCTTTGACCGCATG 1439
Db 487 CCAGCTGATGGGACTAAACGCACTACCATCACTCCGGGGCGCTGGAACCTTAAACCAATG 546
Qy 1440 TTGCGCGCAGCAGAGAGATTTCTATGATGTGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCGCTGCTCTGAAGAAATATGCCATGAATTTGGGCTATTTGGGTAAAGGGAATGTGCT 606
Qy 1500 AGCAAAAAACAATTTGTAGAACAAAGTAGAAGCGGGCGATTTGGTTTAAATTTGCATGAA 1559
Db 607 TATGAACCTCTCCTGCTGATCAACTCGAAGCTGGAGCCATTTGGCTTTAAATTCACGAA 666
Qy 1560 GACTGGGCGACAACACCAAGTGCATCTGCTTGAAGCTGGCAGATGAATACGAT 1619
Db 667 GACTGGGCTAGCACACCTGAGCCATCTACCATTTGCTTGAATGTGGCTGACAAATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGGCGAGGTATGTAGATGACACCTA 1679
Db 727 GTGCAAGTGGCTATCCACACCGATACCTTGAATGAGCGGGCTGTGTGGAAGACACTTTG 786
Qy 1680 AATGCAATGAACGGCGGCCCATCCATGCTTACCACATGAGGAGCGGTGGAGGACAC 1739
Db 787 CAAGCCATTTGCTGGGCGCATTTCCACATTTCCACATGAAGTGTCTGGTGGCGGGAC 846
Qy 1740 TCACCTGATTTATCACCATGGCAGCGAGCTCAATATTCTACCTCTCTCCACACCCCC 1799
Db 847 GCTCCGGATGTCATTAAGATGCTGGCGAATTAACATCTCCAGCTTCTACCAACCCC 906
Qy 1800 ACTATTCCCTATACCAATTAATACGTTGCGAGAACCTTAGACATGCTCATGATGCCAC 1859
Db 907 ACCATTCTCTTCCCGTGAATACAGAAGCGCAACACATGGACATGTTGATGGTGGCCAC 966
Qy 1860 CACCTAGACAAACGGATCCGCGAGGATTTACAATTTCTCAAGCCGATCCGCCCGGC 1919
Db 967 CACTTGGATTAATAACATCAAGAAGATGTCCAGTTTGTGCTGATTTAGGATTCGCCCCAA 1026
Qy 1920 TCTATCGCGCTGAAGATGCTCTCATGATATGCGGTGTGATCCGATGACAGCTCGGAT 1979
Db 1027 ACCATTCGCGCTGAGGACAAATCCACGATATGGGGATTTTCTATATCCAGCTCTGAC 1086
Qy 1980 TCGAAGCAATGGGCGTGCAGCGGAAGTGAATTCCTCGAATTTGGCAGACTGCGGATAAG 2039
Db 1087 TCCCAAGCATGGGCGGTGAGCGAGGTATCACCCTGATCATCCCGGCTTTGGCAACAGCGGACAA 1146
Qy 2040 AATAAAGAAATTTGGTAAAGCTTCTGAAATGGCAAGATGAACATATTTCCGCAAT 2099
Db 1147 AACAAAAAGAAATTTGGTTCGCTTGAGGAAAAAGGGGATAATGACAACTTCCGCATC 1206
Qy 2100 AAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACCCACCGCTGAGCGAGTAT 2159
Db 1207 AAGCGCTACATTTCCAAATACACCATCAACCCCGCTATTTGCAACCGGCAATTTCTGAATAT 1266
Qy 2160 ATCGGCTCTGTGGAAGAGGCGAAGATCGCGCACTTGGTGGTGGGAATCTGCTTTTTTT 2219
Db 1267 GTCGCTCTGTAGAGTGGGCAAAATTCGCGGATTTGGTGTCTTTGGAGTCTCTGCTTTCTTT 1326

Qy 2220 GCGTAAAAACCCAAATTCGTGATCAAAGGCGGTATGCTGTCTTCTCTGAAATGGCGAT 2279
Db 1327 GGCATTAACCCCAACATGATCATCAAAGGCGGATTCATCGCACTTTCTCAAATGGCGAT 1386
Qy 2280 TCTAACGCGTCTGTGCCACTCCCAACCGGTTTATTTACCGCGAAATGTTTGGGCACTAC 2339
Db 1387 GCGAATGSCCTCTATCCCCCACTCCCAACCGGTTGTTTACCGCGAAATGTTTGGGCACTAC 1446
Qy 2340 GCGAAGCGGAAATTTGACACACGATCATCTTTGTTTCCAAAGTCGCTATGAAAAATGCG 2399
Db 1447 GGTAAAGCCAAATTTGACACCAATATCATTTTGTATCCCAAGTGGCTTATGACACGCG 1506
Qy 2400 GTGAAAGAAAAGCTGGGCTTTAGAGCGCAAGTTCTACCGGTCAAAAACTGCCGTAACATC 2459
Db 1507 ATTAAGAAAGAGTTGGGCTTGCAGAGAGTGTTTGTCCAGTTAAAAAATGCGCGCAACATC 1566
Qy 2460 ACCAAGAAAGACTTCAAGTTCAAGCAAAAACGCAAAAATCACCGTCGATCCGAAACACC 2519
Db 1567 ACCAAAAAGACCTCAAAATTCAGATGTTTACCGCACACATCGAAGTCAATCTCTGAAACC 1626
Qy 2520 TTGAGGCTCTTTGATAGTGGCAAACTCTGCACCTCTTAAACCCACTCGCAAGTGCCTCTA 2579
Db 1627 TACAAAGTTAAAGTGGATGGCAAGAGGTTACTTTCCAAAGCGCGATAAAAATCAGCCTA 1686
Qy 2580 GCCAGCGCTACACTTTTCTTTCTAG 2603
Db 1687 GCACAACTCTACAACTTTGTTCTAG 1710

RESULT 8
US-10-476-313-10
; Sequence 10, Application US/10476313
; Publication No. US20040241175A1
; GENERAL INFORMATION:
; APPLICANT: MCKENZIE, BRENT
; APPLICANT: BOYLE, JEFFEREY
; APPLICANT: LEW, ANDREW
; TITLE OF INVENTION: Antigen Targeting
; FILE REFERENCE: BDWP-002
; CURRENT APPLICATION NUMBER: US/10/476,313
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PR5241
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Helicobacter felis
US-10-476-313-10

Query Match 31.3%; Score 902.4; DB 8; Length 1719;
Best Local Similarity 70.6%; Pred. No. 1.1e-195;
Matches 1203; Conservative 0; Mismatches 501; Indels 0; Gaps 0;

Qy 900 AATGTAATAAACAAGNATATGTAATAACCTTACCGACCCCAACCAAGGCGATAAGTGGCG 959
Db 15 AAGATTTACGAAAGAATAATGTTTCTATGTTGTTCCCACTACCGGGGATCGTGTAGA 74
Qy 960 TTAGGAGATACCGATCTTTGGGCAAGAAGTAGAACATGACTATACCACCTATGGCGAAGAA 1019
Db 75 CTCGGGACACTGATTTGATCTTGAAGTGGAGCATGATTCACCACTTATGGTGAAGAG 134
Qy 1020 CTTTAAATTTTGGCGGGTAAAACTATCCGTGAGGGTATGGGTGAGAGCAATAGCCCTGAT 1079
Db 135 ATCAAATTTTGGGCGGTAATACTATCCGTGATGGATGAGTCAAAACCAATAGCCCTAGC 194
Qy 1080 GAAAAACACCTAGATTTAGTTCATCACTAAACGAGTATTCGACTACACCGGATTTTAC 1139
Db 195 TCTTATGAATTTAGATTTGGTGTCTCACTAAACGCCCTCATTTGTGGACTATACGGGCAATTTAC 254
Qy 1140 AAGCCGACATTGGGATTTAAAAACGGCAAAATCCATGCGATTGGCAAGCAGGAACAAAG 1199

Db 255 |||||AAAGCCGACATGGGATTAAGACGCGAAGATTGCGGCATTTGGCAAGGAGGCAATAAG 314
Qy 1200 GACATCAAGATGGCGTAAAGCCCTCATATGGTGTGGGTGGGACAGAGAGCACTAGCA 1259
Db 315 GACATCAAGATGGCGTAGATATAATCTTTGGTAGGTCTGCTACAGAGGCTTTGGCA 374
Qy 1260 GGGAAAGGTATGATTAATACCGCTGGGGGAATCGATTCAACACCCATCTTCCTTTCCCA 1319
Db 375 GCTGAGGGCTTGATTGTAACCGCTGGTGCGATCGATACGCATATTCACTTTATCTCTCCC 434
Qy 1320 CAACAATTCCTTACCGCTTAGCCATGGCTTTACACCATGTTTGGAGCGGCACAGGT 1379
Db 435 CAACAATTCCTTACTGCTTTGCGAGCGGGTTACACCATGATTGGAGGAGGCACAGGA 494
Qy 1380 CTTGTAGATGGCAGAAATCGACTACTATCACTCCGGGCAAAATGGAATTTGCACCGCATG 1439
Db 495 CTTGCGGATGGCAGAAATCGACCAACATCACTCCGAGCGGCTAATCTAAAGATATG 554
Qy 1440 TTGCGCGCAGCAGAGATATTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 555 TTGCGTGCAGCGAAGATACGCATGAATCTAGGCTTTTGGCTAAGGGGAATGTGCT 614
Qy 1500 AGCAAAAAACAATTGTAGAACAAAGTAGAAGCGGGCGCATGTTTTTAAATTTGCATGAA 1559
Db 615 TACGAACCTCTTTTACGCGATCAGATTGAAGCAGGGGCGATTGGTTTTTAAATCCACGAA 674
Qy 1560 GACTGGGCGACACACCAAGTCGATCGATCACTGCTTGAGCGTGGCAGATGAATACGAT 1619
Db 675 GACTGGGGAAGCACACCTCGAGCTATTCACCATCGCTCAATGTGCGCGATGAATACGAT 734
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGCGAGGTATGATGATGACACCTTA 1679
Db 735 GTGCAAGTTGCTATCCACACCGATACCTTAAAGGCGGCTGTGTAAGACACCTTA 794
Qy 1680 AATGCAATGAACGGGCGCGCATTCATGCTTACCAATGAGGAGCGGTGGAGGACAC 1739
Db 795 GAGGCGATTGCGGGCGCACCATCCATACCTTCCACACTGAAGGGGTGGGGTGGACAC 854
Qy 1740 TCACCTGATGTTATCAACATGGCAGGAGCTCAATATCTACCTCTCCACACCCCC 1799
Db 855 GCTCCAGATGTTATCAAAATGGCAGGGGAATTTAAACATTTCTACCCGCTCTACTAACCCG 914
Qy 1800 ACTATTCCTATACCATTAATACGTTGCAAGACACTTACACATGCTCATGACATGCCAC 1859
Db 915 ACAATTCCTTTACCAAAAAACACTGAAGCCGAGCACATGACATGTTAATGGTGGCCAC 974
Qy 1860 CACCTAGACAAACGCGATCCGCGAGGATTTACAATTTTCTCAAGCCGATTCGCGCCCGGC 1919
Db 975 CACTTGGATAAAGTATCAAGGAAGATGTCAGTTTGGCGATTCGAGGATTCGCCCCAA 1034
Qy 1920 TCTATCGCGCTGAAGATGTCTCCATGATATGGGTGTGATTCGGATGACAAGCTCGGAT 1979
Db 1035 ACTATCGCGCTGAAGACCAACTCCATGACATGGGATCTTTTCTATCACCAGCTCCGAC 1094
Qy 1980 TGCAGCAATGGGGCTGAGCGGAGTCAATCTCGAATTTGGCAGACTGCGGTAAG 2039
Db 1095 TCTCAGGCTATGGGACGCGTAGCGAGGTGATCACGCGACTTGGCAGACAGACAAA 1154
Qy 2040 AATAAAAAAGAAATTTGGTAAGTTCCTGAAGATGGCAAGATTAACGATAATTTCCGCATT 2099
Db 1155 AACAAAAAGATTTGGGCGCTTGAAGAGAAAAAGCGGATTAACGAACTTCGCGATC 1214
Qy 2100 AAGCGCTACATCTCAAAATACATATCAACCCGCTTTGACCCACGCGGTGAGCGAGTAT 2159
Db 1215 AAACGCTACATCTTAATACACCATCAACCCCGGATCGCGATGGGATTTCTGACTAT 1274
Qy 2160 ATCGGCTCTGTGAAGAGGCAAGATCGCCGACTTGGGTGTGTAATCTCGCTTTTTT 2219
Db 1275 GTGGGCTCTGTGAAGTGGGCAATACGCCGACCTCGTGTCTTTGGAGTCCGGCTTTCTTT 1334
Qy 2220 GCGTAAACCCCAAAATCGTGATCAAGCGGATGTTGGTCTCTCTGAAATGGGCGAT 2279
|||

Db 1335 GGCATTAAAGCCCAATATGATTATTAAAGGCGGATTTATTTCGCTCTCTCAAATGGCGAT 1394
Qy 2280 TCTAACCGCTGTGTCGCCACTCCCAACACCGGTTTATTATCCGCGAAATGTTTGGGATCAC 2339
Db 1395 GCCAATCGCTATTTCACCCCTCAGCCCGCTATTATCCCGTGAATGTTTGGACACCAT 1454
Qy 2340 GGCAGGCGAAATTTGACACGAGCATCACTTTGTTCCAAAGTCGCTTATGAAATGCG 2399
Db 1455 GGGAAAAACAATTCGACACCAATATCACTTTCGTGTCCCAAGCGGCTTACAAGGCGGG 1514
Qy 2400 GTGAAAGAAAAAGCTGGGCTTAGAGCGCCAAAGTTCTTACCGTCAAAAACTGCCGTAAATC 2459
Db 1515 ATCAAGAGAACTAGGCTAGATCGGTGTATTTGCCAGTGAATACTGTGCAATATC 1574
Qy 2460 ACCAAGAAAGACTTTCAAGTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAAAAC 2519
Db 1575 ACTAAAAAGGACCTCAAAATTCACGATGTGACCGCACATATTGATGTCAACCTGAAAC 1634
Qy 2520 TTCGAGTCTTTGTAGATGGCAAACTGTGACCTCTTAAACCCACCTCGCAAGTGCCTTA 2579
Db 1635 TATAGGTGAAGTGGATGGCAAGAGGTAACCTCTAAAGCAGCAGATGAATTGAGCCTTA 1694
Qy 2580 GCCAGCGCTACACTTTCTCTCTAG 2603
Db 1695 GCGCAACTTTATAATTTGTTCTAG 1718

RESULT 9

US-10-335-977-3849
; Sequence 3849, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragourae, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3849:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori

FEATURE:									
NAME/KEY: misc feature									
LOCATION: (B) LOCATION 1...1710									
SEQUENCE DESCRIPTION: SEQ ID NO: 3849:									
US-10-335-977-3849									
Query Match 30.7%; Score 886.4; DB 7; Length 1710;									
Best Local Similarity 70.0%; Pred. No. 5.1e-192;									
Matches 1193; Conservative 0; Mismatches 511; Indels 0; Gaps 0;									
QY	900	AAAAAGAAAAACAAGAAATATGTAATACCTAGGACCCACCAAAAGCGGATAAAGTGGCG	959						
DB	7	AAGATTAGCAGAAAAGAAATATGTTTCTATGTATGGCCCTACCACGGCGATAAAGTGAGA	66						
QY	960	TTAGGAGATPACCGATCTTTTGGGCAGAGTAGAACATGACTATACCACTATATGGCGGAAGAA	1019						
DB	67	TTGGCGGATACAGATTGTATCGCTGAAGTAGAACATGACTACACCATTTATGGCGAAGAG	126						
QY	1020	CTTAAATTTGGCGCGGGTAAACTATCCGTGAGGGTATAGGGTCAGAGCAATAGCCCTGAT	1079						
DB	127	CTTAAATTCGGTGGCGGTAAAAACCCCTAAGAGAAGGCATGAGCCAATCCAAACCCCTAGC	186						
QY	1080	GAAPACACCTAGATTAGTCACTACCTAAGCGATGATTATCGACTACACCGGGATTAC	1139						
DB	187	AAAGAAGAACTGGATTGTATCATCACTAAGCGTTAATCGTGGATTACACCGGTATTAT	246						
QY	1140	AAAGCCGACATTTGGGATTAAAAACGGCAAAATCCATGGCATTTGGCAAGCGAGGAAAACAAG	1199						
DB	247	AAAGCGGATTTGGTATTAAAGCGGCAAAATCGCTGGCATTTGGTAAAGCGGTTAACAA	306						
QY	1200	GACATGCAAGATGGCGTAAAGCCCTCATATGGTCTGGGTGTGGGCAACAGAACCTAGCA	1259						
DB	307	GACATGCAAGATGGCGTTAAAAACAATCTTAGCGTGGTCTCTACTGAAGCGCTAGCT	366						
QY	1260	GGGGAAGGTATGATTATTACCGCTGGGGATCGATTTCACACCCACTTCCTTTCTCCA	1319						
DB	367	GGTGAAGGCTTAATCGTAAACGGCTGGTGTATTGACACACATCCACTTTCATTTACCC	426						
QY	1320	CAACAATTCCTTACCGCTCTAGCCAATGGCTTACAAACCATGTTTGGAGCGGCACAGGT	1379						
DB	427	CAACAATTCCTTACAGCTTTTGCAGCGGTGTAAACCAACATGATTGTGGCGGAACCGGT	486						
QY	1380	CCTGTAGATGGCAATGGCACTACTATCACTCCGGGCAAAATGGAACTTGCACCGCATG	1439						
DB	487	CCTGCTGATGGCACTAATCGCACTACTATCACTCCAGGTAGAGAACAATTTAAATGGATG	546						
QY	1440	TTGGCGGACAGAGAAGATATTCTTATGAATGGGCTTTTGGGCAAAAGGCAATAGCTCT	1499						
DB	547	CTCAGAGCGGCTGAAGAAATATTCTATGAACCTTAGGTTTCTTAGCTAAAGGTTAACGGTTCT	606						
QY	1500	AGCAAAAACAACCTTGTAGAACAGTAGAAGCGGCGCATTTGGTTTAAATTTGCATGAA	1559						
DB	607	AACGACGCGAGCTTAGCGGATCAAAATGAAGCGGTCGGAATGGCTTTTAAATTTCCAGAA	666						
QY	1560	GACTGGGGCACAAACCAAGTCGATCGATCACTGTCTTGAGCGTGGCAGATGAATACGAT	1619						
DB	667	GACTGGGGCACCACTCTTCTGCAATCAATCATGCGTTAGACGTTGGCGCAATACGAT	726						
QY	1620	GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGTTATGTPAGATGACACCTTA	1679						
DB	727	GTGCAAGTCGCTATCCACACAGACACTTTGAATGAAGCCGGTTGCGTGGGAAGACACTATG	786						
QY	1680	AATGCAATGAACGGGGCGCCATCCATGCCATACCAATTGAGGAGCGGTTGGAGGACAC	1739						
DB	787	GCAGCTATTTCGCGGACGCAATATGCAACTTTCCACACTGAAGCGCTGGTGGCGGACAC	846						
QY	1740	TCACTGTATGTTATCACCTGGGAGCGAGCTCAATATTCTACCTCTCCACACACCC	1799						
DB	847	GCTCCTGACATATTAAAGTGGCGGTGAACACACATCTACCCGCTTCCACTAACCC	906						
QY	1800	ACTATTCCCTATACCAATTAATACGGTTGCAAGACACTTAGACATGCTCATGACATGCCAC	1859						
DB	907	ACTATCCCTTTCACTGTGAATACAGAAGCGGCAACACATGACATGCTCATGGTGTGCCAC	966						

QY	1860	CACCTAGACAAACGCAATCCGCGAGGATTTACAAATTTTCTCAAAGCCGTTATCGGCCCGGC	1919						
DB	967	CACCTGGATAAAAGCAATTAAGAAGATGTCCAGTTCCGTGATTCAAGGATCCGCCCTCAA	1026						
QY	1920	TCTATCGCGGCTGAAGATGTCTCCATGATATGGTGTGATCGCGATGACAAGCTCGAT	1979						
DB	1027	ACCAITTCGCGCTGAAGACACTTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC	1086						
QY	1980	TCCCAAGCAATGGGCGTCAGGCGGAAGTGAATCTCTGAACCTTGGCAGACTGCGGATAG	2039						
DB	1087	TCTCAAGCGATGGTGTGTGGTGAAGTTATCACTAGGACTTGGCAAAACAGCTGACAA	1146						
QY	2040	AATAAAAAAGAAATTTGGTAAAGCTTCTTGAAGATGGCAAGATAACGATAAATTTCCGCATT	2099						
DB	1147	AACAAAAAGAAATTTGGCCGCTTGAAGAAGAAAAAGGCGATAACGACAACTTCAGGATC	1206						
QY	2100	AAGCGCTACATCTCCAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT	2159						
DB	1207	AAAACGCTACTTGTCTAAATACACCAATTAACCCAGCGATCGCTCATGGGATTTAGCGAGTAT	1266						
QY	2160	ATCGGCTCTGTGAAGAGGCAAGATCGCGACTTGGTGGTGTGGAATCCCTGCCCTTTT	2219						
DB	1267	GTAGTTCTGTAGAAGTGGGCAAGTGCTGACTTGGTATTGTGGAGTCCAGCATTCITT	1326						
QY	2220	GGCGTAAAAACCAAAATCGTGATCAAAAGGCGGTATGGTGGTCTTCTCTGAAATGGGCGAT	2279						
DB	1327	GGCGTGAACCCCAACATGATCATCAAGGCGGATTCATTGCGTTAAGCCAAATGGGCGAT	1386						
QY	2280	TCTAACCGCTGTGTGCCACTCCCAACCGGTTTATTACCGCGAAATGTTTGGGGATCAC	2339						
DB	1387	GCGAACGCTTCTATCCCTACCCCAACACCGGTTTATTACAGAGAAATGTCGCTCACCAT	1446						
QY	2340	GGCAAGCGAAATTTGACACGACATCACTTTTGTTCCAAAGTCGCTATGAAATGCG	2399						
DB	1447	GGTAAAGCTAAATACGATCAAAACATCACTTTTGTGCTCAAGCGGCTTATGACAAAGC	1506						
QY	2400	GTGAAGAAAAAGCTGGGCTTAGAGCGCAAGTCTTACCGGTCAAAAACTGCGGTAAACATC	2459						
DB	1507	ATTAAAGAAGNATTAGGCTTGAAGACAGTGTTCGCGTTAAAAATTCGAGAAACATC	1566						
QY	2460	ACCAAGAAAGACTTCAAGTTCAACAGCAAAAACGCAAAAATCAACGTCGATCCGAAAACC	2519						
DB	1567	ACTAAAAAGACATGCAATTCATGACACTACCGCTCAATTTGAAGTCAATCCTGAAACT	1626						
QY	2520	TTGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTAAACCCACTCCGNAAGTGCCTCTA	2579						
DB	1627	TACCATGTGTCTGGATGGCAAGAAAGTCACTTCTTAAACCAAGCTAATAAAGTGAGCTTG	1686						
QY	2580	GCCGAGCGCTACACTTTTCTCTAG	2603						
DB	1687	GCTCAACTCTTTAGCATTTTCTAG	1710						

RESULT 10

US-09-895-913A-251
; Sequence 251, Application US/09895913A
; Patent No. US20020160456A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Omen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 251

; LENGTH: 1815

; TYPE: DNA

; ORGANISM: Helicobacter pylori

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (51)...(1757)

US-09-895-913A-251

Query Match 30.7%; Score 885.8; DB 3; Length 1815;

Best Local Similarity 70.0%; Pred. No. 7.2e-192;

Matches 1193; Conservative 0; Mismatches 512; Indels 0; Gaps 0;

Qy	900	AAAAAGAAAAACAAGAAATATGTAATACCTACGGACCCACCAAGGCGATAAAGTGCGC	959
Db	57	AGATTAGCAGAAAGAAATATGTTCTATGTATGGCCCTACTACAGCGGTAAGTGAGA	116
Qy	960	TTAGGAGATACCGATCTTTGGGAGAGTAGAAATGACATGATATACCATGATGGCGAAGAA	1019
Db	117	TTGGCGATACAGACTTGATCGCTGAAGTAGAATGACATGACATACCATTTATGGCGAAGAG	176
Qy	1020	CTTAATTTGGCGCGGTAATACTATCGTAGGGATATGGGTGAGAGCAATAGCCCTGAT	1079
Db	177	CTTAATTCGGTGGCGGTAAACCCCTGAGAGAGGATGAGCCAATCCAAACCCCTAGC	236
Qy	1080	GAACACACCTAGATTTAGTATCATCACTAAGCGATGATTTATGACATACACCGGATTTAC	1139
Db	237	AAAGAAGAAATTTGATCTAATCATCACTAAGCGTTTAATCTGGATTTACACCGGTATTTAT	296
Qy	1140	AAAGCGCATTTGGGATTTAAACAGGCAAAATCCATGGCATTTGGCAAGGCGAGAAACAAG	1199
Db	297	AAAGCGGATTTGGTATTAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGTACAAA	356
Qy	1200	GACATGCAAGATGGCGTAAGCCCTCATATGGTGGTGGTGGGCAAGAGCACTAGCA	1259
Db	357	GACATGCAAGATGGCGTTAAACAAATCTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCC	416
Qy	1260	GGGGAAGGTATGATTTATCCGCTGGGGAATCGATTCACACACCCACTTCTTTCTCCA	1319
Db	417	GGTGAAGGTTTGATCGTAACTGCTGGTGGTATTTGACACACATCCACTTCAATTTACCC	476
Qy	1320	CAACAAATCCCTACCGCTCTAGCAATGGCTTACAAACCATGTTTGGAGCGGCACAGGT	1379
Db	477	CAACAAATCCCTACAGCTTTGCAAGCGGTGTAACCAACATGTTGGTGGCGAATCGGT	536
Qy	1380	CCTGTAGATGGCAAGATCGAATCTACTATCACTCCGGGCAAAATGGAATTTGCAACGATG	1439
Db	537	CCTGTGTATGGCACTAAATGCGACTACTATCACTCCAGGCGAAGAAATTTAAATGGATG	596
Qy	1440	TTGGCGGAGCAGAGAGATTTCTATGATGTGGGCTTTTGGGCAAGAGCAATAGCTCT	1499
Db	597	CTCAGAGCGGCTGAAGAAATTTCTATGAATTTAGGTTTCTTGGCTAAAGGTAAACGCTTCT	656
Qy	1500	AGCAAAACAACTGTGTAGAAAGTAGAAGCGGCGCATTTGGTTTAAATTTGCAATGAA	1559
Db	657	AACGCGCAGCTTAGCCGATCAATTTGAAGCTGGTGGCATTTTAAATTTCCAGGAA	716
Qy	1560	GACTGGGCGACACACCAAGTGGATCGATCTGCTTGGCGTGGCGAGATGAATACGAT	1619
Db	717	GACTGGGCGACCACTCTTCTGCAATCAATCATGCTGCTAGATGTTGCAGACAAATACGAT	776
Qy	1620	GTGCAAGTTTGTATCCACACCGATACAGTCAATGAGCGAGGTATGTAATGACACCCCTA	1679
Db	777	GTGCAAGTTCGTATCCACACAGACACTTTGAATGAAGCGGTTGCGTGGGAAGACACTATG	836
Qy	1680	AATGCATGAACGGGCGGCCATCCATGCTTACCACTTGGAGGAGCGGTGGAGGACAC	1739
Db	837	GCAGCTATTGGCGGAGCGCACTATGCACTTTCCACACTGAAGGTGCTGGCGCGGACAC	896
Qy	1740	TCACCTGATGTTATACCATGGCAGCGAGCTCAATATTTCTACCTCTCCACACCCCC	1799
Db	897	GCTCTGATATTTAAAGTAGCTGGTGAACACACATTTCTTCCCGTTCCACTAACCC	956

Qy	1800	ACTATTCCTTATACCAATTAATAACGGTTGCAGAACACTTAGACATGCTCATGACATGCCAC	1859
Db	957	ACTATCCCTTTCACTGTGAATACAGAAAGACACACATGGACATGCTTTATGGTGTGCCAC	1016
Qy	1860	CACCTAGACAAAACGATCCGGAGAGATTTCATTTTCTCAAGCCGATTCGCCCGCGC	1919
Db	1017	CACCTGGATAAAGCATTAAGAAAGATGTTTCAGTTTCGCTGATTTCAAGGATTCGCCCTCAA	1076
Qy	1920	TCATTCGGCTGAAGATGCTCATGATATGGTGTGATCGCGATGACAAAGCTCGGAT	1979
Db	1077	ACATTCGGCTGAGACACTTTGATGACATGGGATTTTCTCAATCACCAGCTTGAC	1136
Qy	1980	TGCAAGCAATTTGGGCGGTGAGGCGAAAGTGAATCTCGAACTTTGGCAGACTGCGGATAAG	2039
Db	1137	TCTCAAGCTATTTGGTGGTGGTGAAGTTATCACTAGAACTTGGCAAAACAGCTGACAAA	1196
Qy	2040	AATAAAGAAATTTGGTGAAGTCTTCGAGATGCGCAAGATGCAAGATTAATTTCCGCAT	2099
Db	1197	AACAAAAGAAATTTGGCGCTTGAAGAGAAAGGCGATTAACGACAACTTTAGGATC	1256
Qy	2100	AAGCGTACATCTCCAAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGAGTAT	2159
Db	1257	AAAGCTACATTTCTAATACCAATTAACCCAGGATCGCTCATGGATTTAGCGAGTAT	1316
Qy	2160	ATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTTGGTGGTGTGGAATCTTGCCTTTTT	2219
Db	1317	GTAGGTTCTGTAGAGTTGGCAAGTGGCTGACTTGGTATTTGTTGGAGTCCCGCATTTCT	1376
Qy	2220	GGCGTAAACCCAAATTCGTGATCAAGGCGGTATGGTGGTCTTCTCTGAATTTGGCGAT	2279
Db	1377	GGCGTAAACCCAAATGATCATCAAGGCGGTTTCAATTTGGTGGTGAATTTGAGTCAAT	1436
Qy	2280	TCATAAGCGCTGTGCGCACTCCCAACCGGTTTATTTACCGGAAATGTTTGGGATCAC	2339
Db	1437	GCAGAACGCTTCTATCCCTACCCCAACCAAGTTTATTTACAGAGAAATGTTGCTCATCAT	1496
Qy	2340	GGCAAGCGAAATTTGACACCGCATCACTTTTGTTCCTTCAAGTCCCTATGAAATGGC	2399
Db	1497	GGTAAAGCAATACGATGCAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGC	1556
Qy	2400	GTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAACTGCGGTAAACATC	2459
Db	1557	ATTAAGAGAAATTTAGGCTTGAAGAGCAAGTTTGGCGTTAAAAATTTGCAAGAAATC	1616
Qy	2460	ACCAAGAAAGCTTCAAGTTCAACGCAAAACGCAAAATCAACGTCGATCCGAAAC	2519
Db	1617	ACTAAAAAGACATGCAATTTCAACGCACTAACCGCTCACTTGAAGTCAAATCTGAAACT	1676
Qy	2520	TTGAGGCTTTGTAGATGGCAACTCTGCACTCTTAAACCCACCTCGCAAGTGCCTCTA	2579
Db	1677	TACATGTTGTTGGATGGCAAGAGTAACTTTAAACCAACCAATTAAGTGAAGTTG	1736
Qy	2580	GCCGAGCGCTACACTTTTCTTAGG	2604
Db	1737	GCGCACTCTTTAGCATTTTCTAGG	1761

RESULT 11

US-10-282-122A-22427

; Sequence 22427, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Hasebeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

Qy 2520 TTCGAGGTCTTTGTAGATGCCAAACTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTA 2579
Db 1627 TACCATGTGTCTGTGGATGCCAAAGAGTAACCTCTAAACCCAGCAATAAAGTGAGCTTG 1686
Qy 2580 GCCAGGCTACACTTCTTCTTAG 2603
Db 1687 GCGCAACTCTTTAGCAATTTTCTAG 1710

RESULT 12

US-10-500-447A-5
; Sequence S, Application US/10500447A
; Publication No. US20050150016A1
; GENERAL INFORMATION:
; APPLICANT: PARK, Hee-Sung
; TITLE OF INVENTION: Method for producing a recombinant protein using pollen
; FILE REFERENCE: YIOP040518US/PCT
; CURRENT APPLICATION NUMBER: US/10/500,447A
; CURRENT FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: KR 2001-71712
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-500-447A-5

Query Match 30.5%; Score 880; DB 9; Length 1710;
Best Local Similarity 69.8%; Pred. No. 1.5e+190;
Matches 1189; Conservative 0; Mismatches 515; Indels 0; Gaps 0;

Qy 900 AAAATGAAAAACAAGATATGTAATACCTACGAGCCACCAAGGCGATAAAGTCCG 959
Db 7 AAGATTAGCAGAAAAAGAAATATGTTCTATGTATGGCCCTACTACAGGCGATAAAGTGAGA 66
Qy 960 TTAGGAGATACCGATCTTTGGGAGAGTAGACATGACTATACCCCTATGGCGAGAA 1019
Db 67 TTGGGCGGATACAGACTTGATCGTGAAGTAGAGACATGACTACACCACTTTATGGCGAAGAG 126
Qy 1020 CTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGTACAGACCAATAGCCCTGAT 1079
Db 127 CTTAAATTTGGCGGGTAAACCTATCCGTGAGGGTATGGGTACAGACCAATAGCCCTGAT 186
Qy 1080 GAAACACCCCTAGATTAGTATCATCACTAACCGGATGATTATCGACTACACCGGATTTAC 1139
Db 187 AAAGAAGAACTGGATCTAATCATCACTAACCGCTTTAATCGTGGATTACACCGGTATTTAT 246
Qy 1140 AAAGCCGACATTTGGGATTTAAACCGGCAAAATCCATGGCAATTTGGCAAGGAGGAAACAAG 1199
Db 247 AAAGCGGATATTGGTATTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGGTAAACAAA 306
Qy 1200 GACATCGAGATGGCTTAAGCCCTCATATGGTGGTGGTGGGACAGAGCACTAGCA 1259
Db 307 GACATCGAAGTGGCTTAAACCAATCTTAGCTGGGGTCTGCTACTGAAGCCCTTAGCC 366
Qy 1260 GGGGAAGGTATGATTAATTAACCGTGGGGGAATCGATTACACACCCCACTTCCTTTCTCCA 1319
Db 367 GGTGAAGTTTGTATGTAATCTGTTGGTATTGACACACATCCATCTCATCTCCCC 426
Qy 1320 CAACAAATTCCTACCGCTTAGCCAAATGGCGTTACAAACCATTTTGGAGCGGCAAGGT 1379
Db 427 CAACAAATTCCTACAGCTTTGCAAGCGGTAAACACCATGATTGGTGGCGGAATCGGC 486
Qy 1380 CCTGTAGATGGCAGATGCGACTACTATCACTCCGGGCAATGGAATTTGCACCGCATG 1439
Db 487 CCTGTGATGGCACTAACGCAACCACTATCACTCCAGGTAGAGAAATTTTAAATTTGGAT 546
Qy 1440 TTCCGCGCAGCAGAGATTTCTATGAATGGGGCTTTTGGGCAAGGCAATAGCTCT 1499
Db 547 CTCAGAGCGCAGAGATATTTCTATGAACCTTAAGTTTCTTAGCTAAAGGTAAAGCTTCT 606

Qy 1500 AGCAAAAAACAACCTTGTAGAAACAAGTAGAGCGGCGCATTTGGTTTAAATTTGCATGAA 1559
Db 607 AACGATGCAAGCTTTAGCCGATCAAATTTGAAGCGGTGCGATTTGGCTTTAAAAATCCACGAA 666
Qy 1560 GACTGGGGCAACAACCAAGTGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGAT 1619
Db 667 GACTGGGGCACCACTCTCTTCTGCAATCAATCATGCGTTAGATTTGCGGACAAATACGAT 726
Qy 1620 GTGCAAGTTTGTATCCACACCGCATACAGTCAATGAGCGAGGTTATGTAGATGACACCCCTA 1679
Db 727 GTGCAAGTTCGCTATCCACACAGACACTTTTGAATGAAGCGGTTGTGTAGAGACACTATG 786
Qy 1680 AATGCAATGAACGGCGCGCCATCCATGCCCTAACCATTTGAGGGAGCGGTTGGAGACAC 1739
Db 787 GCAGCCATTCGCGGACGCACTATGCACACTTTCCACACTGAAGGCGCTGGTGGCGACAC 846
Qy 1740 TCACCTGATTTATCAACATGGCAGGAGCTCAATATTTCTACCTCTCCACCAACCC 1799
Db 847 GCTCCTGATATTTAAAGTAGCTGGTGAACACACACTTTCTGCCGCTTCCACTAACCC 906
Qy 1800 ACTATTCCTATACCATTAATACGTTGCAGAACACTTAGACATGCTCATGACATGCCAC 1859
Db 907 ACTATCCCTTTCACTGTGAATACAGACGACACATGGACATGCTTATGTGTGCCAC 966
Qy 1860 CACCTAGACAAACGCACTCCGCGAGGATTTTCAATTTTCTCAAGCGCTATCCGCCCGGC 1919
Db 967 CACTTGGATAAAGCATTTAAGAGAGATGTTCACTTCGCTGATTCAGGATCCGCCCTCAA 1026
Qy 1920 TCTATCCGCGTGAAGATGTCTCATGATATGGGTGTGATCGCATGACAAAGCTCGGAT 1979
Db 1027 ACTATTCGCGTGAAGACACTTTGCATGACATGGGATTTTCTCAATCACCAGTTCTGAC 1086
Qy 1980 TCGCAAGCAATGGGCGTGCAGGCAAGTGAATCTCGAACTTGCAGACTTGCAGATAG 2039
Db 1087 TCTCAAGCTATGGTCTGTGGTGAAGTTATCACCAGAACTTGGCAACAGCTGACAAA 1146
Qy 2040 AATAAAAAGAAATTTGGTAAGCTTCTCAAGATGGCAAGATAACGATAAATTTCCGCAT 2099
Db 1147 AACAAAAAGAAATTTGGCGCTTGAAGAGAAAGAGCGATTAACGACAACTTCAGGATC 1206
Qy 2100 AAGCGCTACATCTCCTCAATACACTATCAACCCCGCTTTGACCCAGCGGTGAGCGATAT 2159
Db 1207 AAACGCTACTTGTCTAAATACACCATTAACCCAGCGATCGCTCATGGGATTAAGCGATAT 1266
Qy 2160 ATCGGCTCTGTGAAGAGGCGCAAGATCCCGACCTTGGTGGTGGAAATCTCGCTCTTTT 2219
Db 1267 GTAGGTTCTGTAGAAGTGGGCAAGTGGCTGACTTGGTGGTGGAGTCCCGCATTTCTTT 1326
Qy 2220 GCGTAAACCCCAAAATCGTATCAAGCGCGTATGGTGGTCTTCTCTGAATGGCGAT 2279
Db 1327 GCGTGAACCCCAACATGATCATCAAGCGCGATTCATTTGATTTAGTCAAAATGGGTAT 1386
Qy 2280 TCTAACCGCTGTGTGCCCACTCCCAACCGGTTTATTAACCGGAAATGTTTGGGATCAC 2339
Db 1387 GCGAAGCTTCTATCCCTACCCCAACCGGTTTATATAGAGAAATGTTGCTCATCAT 1446
Qy 2340 GGCAAGCGAAATTTGACACCAAGCATCACTTTTGTTCCAAAGTGCCTATGAAATGGC 2399
Db 1447 GGTAAAGCTAAATACGATGCAAAACATCACTTTTGTCTCAAGCGGCTTATGACAAAGC 1506
Qy 2400 GTGAAGAAAGCTGGGCTTAGAGCGCAAGTTCTACCGGTCAAAAACCTGCGGTAACTC 2459
Db 1507 ATTAAAGAAAGTATTAGGCTTTGAAGGCAAGTGTGGCGGTAAAAAATTTGCAAAACATC 1566
Qy 2460 ACCAAGAAAGACTTTCAAGTTCAACGACAAAACGCAAAAATCACCGTGCATCCGAAAAC 2519
Db 1567 ACTAAAAAGACATGCAATTCACGACACTACCGTCACTTGAAGTCAATCTGAACT 1626
Qy 2520 TTCGAGTCTTTGTAGTGGCAAACTCTGCACTCTTAAACCACTCGCAAGTGCCTCTA 2579
Db 1627 TACCATGTGTTGTTGGATGGCAAGAGTAACCTTCTAAACCAAGCAATAAAGTGAGCTTG 1686

Qy	1524	GTAGAAAGCGCGCGATGGTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGGC	1583
Db	1259	ATAACAGCGGGTGTATAGTCTTAAATATACATGAAGACTGGGGGCAACGCCAATGGCA	1318
Qy	1584	ATCGATCACTGTTGAGCGTGGCAGATGAATAGATGTGCAAGTTTGTATCCACCGAT	1643
Db	1319	ATTCAAAATTGGCTTAAATGTCGCGGATGAATGGATGTACAGGTGGCTATTCATCTGAC	1378
Qy	1644	ACAGTCAATGAGCGAGGTTATGTAGATGACACCTTAAATGCAATGAACGGCGCGGCATC	1703
Db	1379	ACCTTAAATGAAGTGGTTTTATGAAGACAGTAAAGCCATTCGCCGTGAGTGATC	1438
Qy	1704	CATGCCCTACCAATTTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGGCA	1763
Db	1439	CATGTATTCATACCGAAGCGCAGGTGGTGGTTCATGCCCTCATGTGTGATCAAGTCGGTA	1498
Qy	1764	GGGAGCTCAATATTTCTACCTCTCCCAACCCGCCACTATTCCTTATACCATTAATACG	1823
Db	1499	GGAGAGCCCAATATTTTACCTGCATCAACCAACCAACGATGCTTATACCATTAATACC	1558
Qy	1824	GTTCAGACACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGGAG	1883
Db	1559	GTGGACGAGCATCTTGATATGTTGATGTCATCATCTCGATCCCTCATTTCTGAA	1618
Qy	1884	GATTTACAATTTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTCTC	1943
Db	1619	GATGGGCAATTTGCTGAACTCGTATTCGTCGGAACCAATTCGTCGAGNAGATATCTTA	1678
Qy	1944	CATGATATGGGTGTGATCGCGATGACAAGCTCGGATTCGCAACAATGGGGCGTGCAGGC	2003
Db	1679	CATGATATGGGGCAATTTTCGGTGTGTCGTCAAGCTCAACGCCATGGGACGAGTCGGA	1738
Qy	2004	GAAGTGATTCCTGGAATCTGGCAGACTCGGATAGAAATAAAGAAATTTGGTAAGCTT	2063
Db	1739	GAAGTATCTTACGCCACTTTGGCAGTGTGCACATAAAATGAAATTGCAACGAGGCACATTA	1798
Qy	2064	CCTGAGATGGCAAGATTAACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACT	2123
Db	1799	GCGGTGATAGCGCATATGATATAATATCGTATTAACGTTATATCGCTAAATACAG	1858
Qy	2124	ATCAACCCCGCTTTTGACCCACGCGTGGAGTATATCGGCTCTGTGGAAGAGGGCAAG	2183
Db	1859	ATTAAACCCGCACTGCACATGATTTGCTCATACGGTGGATCAATAGAAAAGGTAAA	1918
Qy	2184	ATCGCCGACTTGGTGTGTGGAATCTGCTTTTGGCGTAAAAACCAAAATCGTGATC	2243
Db	1919	CTTGGGATATCGTGTATGGGATCCTGCTTTCTTTGGCGTCAACCGGCATTTATCATTA	1978
Qy	2244	AAAGCGGTATGGTGTCTTCTGAAATGGGCGATTTCAACGGCTCTGTGCCACTCCC	2303
Db	1979	AAAGTGGTATGGTTCGCTTATGCGCAATGGGGGATATTAATCGCGCTATTCACACCG	2038
Qy	2304	CAACCGTTTATTACCGCAAAATGTTTGGGCATCAGCGAAGCGGAAATTTGACCCAGC	2363
Db	2039	CAACCGTTTCAATATCGTCCAATGATGCCCTTTTAGGAAAGCCAAATATCAACGTCG	2098
Qy	2364	ATCACTTTTGTTCCTAAAGTTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTAGAG	2423
Db	2099	ATGATCTTTATGTCAAAAGCGGATTTAGGCGGGAGTGCCAGAAAAATTAGGCTTAAAA	2158
Qy	2424	CGCAAGTTCCTACCGGTCAAAAACCTGCGGTAAATCAATCCAAAGAAAGCTTCAAGTTCAAC	2483
Db	2159	AGCTTAAATGGTGTGGAGGCTGTGCTCATATCACAAAAGCTTTCGATGATCCCAAT	2218
Qy	2484	GACAAAACGGCAAAATCACCGTCCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAA	2543
Db	2219	AACATGTTCTCATATCGAATTAGATCCAAACCTTACATTTTAAAGCGGATGGTGA	2278
Qy	2544	CTCTGCACCTCTAAACCCACCTCGCAAGTGCCTCTAGCCCGAGGCTACACTTTCTCTAG	2603
Db	2279	CCACTGGTTGTGAGCCAGGACTGAATTAACCGATGGCTCAACGCTATTTCTTATTTAA	2338

Qy	2604	GCA	2606
Db	2339	CCA	2341

Search completed: November 29, 2005, 08:04:19
Job time : 2060 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:34:45 ; Search time 563 Seconds
(without alignments)
760.989 Million cell updates/sec

Title: US-09-904-994B-1
Perfect score: 2883
Sequence: 1 rgragattttccaractt.....aaaaagtagagccacagg 2883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3205263 seqs, 74304013 residues

Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Watch 0%
Maximum Watch 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	612	21.2	3164	1	US-10-793-626-3828
2	612	21.2	3234	1	US-10-793-626-3853
3	578.6	20.1	1716	1	US-10-793-626-117
4	455	15.8	1612	1	US-10-485-517-26
5	424.4	14.7	3475	1	US-10-793-626-4032
C 6	216.2	7.5	526	1	US-10-834-151-2
C 7	97.4	3.4	309	1	US-10-793-626-121
8	74.2	2.6	336	1	US-10-793-626-115
9	46	1.6	2259	1	US-10-793-626-4078
C 10	40	1.4	3373	1	US-10-793-626-4389
C 11	40	1.4	3927	1	US-10-793-626-3926
C 12	39.8	1.4	3923	1	US-10-793-626-4109
C 13	37.8	1.3	1809	1	US-10-793-626-2361
C 14	37.8	1.3	3591	1	US-10-793-626-4033
C 15	37.2	1.3	2988	1	US-10-793-626-4132
C 16	37.2	1.3	4210	1	US-10-793-626-3761
C 17	37	1.3	3362	1	US-10-793-626-3731
C 18	37	1.3	4041	1	US-10-793-626-3689
C 19	36.2	1.3	738	1	US-10-793-626-2909
C 20	36.2	1.3	3218	1	US-10-793-626-4229
C 21	36.2	1.3	3997	1	US-10-793-626-4220
C 22	35.8	1.2	14941	1	US-10-821-234-771
C 23	35.6	1.2	1017	1	US-10-793-626-1003

ALIGNMENTS

RESULT 1
US-10-793-626-3828
; Sequence 3828, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P0348005
; CURRENT APPLICATION NUMBER: US/10793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3828
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3828

Query Match	21.2%;	Score 612;	DB 1;	Length 3164;
Best Local Similarity	57.4%;	Pred. No. 5.2e-139;		
Matches 1211;	Conservative 0;	Mismatches 845;	Indels 52;	Gaps 4;
Qy	543	TTGGTTCGATAAGACATCGAGCTCAATCGAGGCAAGAGTAAACCGAACTTTGAGGTTA	602	Sequence 41, Appl
Db	468	TTGTTAAATAACTGAAATAGAGTCAATAACATCATCCCGAAACGGTGTAGTAA	527	Sequence 56, Appl
Qy	603	CTAATGAAGGCGCTAAATTCCTTGCATGTGGGTAGCCATTTCCATCTTTTGAAGTAAACA	662	Sequence 3694, Ap
Db	528	AAATACAGGCGATAGACCTATACAAGTAGGTTCACATTTCCACCTTTTGAAGCAATA	587	Sequence 4222, Ap
Qy	663	AGGCATTAATTCGATCGTGAAGAAAGCCTATGGCAACCGCTAGATATTCCTCTGGCA	722	Sequence 4141, Ap
Db	588	AAGCATTAGATTTGATCGTGAAGAAAGCATATGGTAAACATTTGGATATTCCTGCGAGAG	647	Sequence 3, Appli
Qy	723	ACACGCTACGATTTGGGCGAGCAACCCCGAAGTCAGTTGATTCCTCTTTGGTGCA	782	Sequence 3919, Ap
Db	648	CTGCAGTGAGATTTGAACCTCGAGATGAAAAAAGTAAACACTTGTGATATTCGAGAC	707	Sequence 3885, Ap
Qy	783	GTAAGAAAGTATTTGGATGATGAAACGGGCTGTGTAATAAC- --ATCGCGGATGAACGCATA	839	Sequence 61, Appl
Db	708	GACGTAAATTTATGGATTCCGTGGTTAGTCGATGGCGATATTGACGAAGAACGCGTAT	767	Sequence 3343, Ap

Qy	840	AACATAAAGCGCTTGACAAGCGGAAATCTCACGGATTTTATCAAGTAAGGAGACTCCCATG	899
Db	768	TCGGTCCAAATGATTCAAAATCAAAAAGCCGCCGTTAAAAACGATGAGGCGAAGACAATG	827
Qy	900	-----AAAATGAAAAACAAGATATGTAA	925
Db	828	CGAATAAAAAGGTGGTAAATAGAGTAGATTTTAAATGACAACTCTCAATACACAAG	887
Qy	926	TACCTACGGACCCACAAAGCGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCAGA	985
Db	888	TCTTTATGGACCAACTGTAGGAGACTCTGTGAGATTAGGAGATACGAACTGTTTTCACA	947
Qy	986	AGTAGAACATGACTATACACCTATGTCGGAAGAACTTTAAATTTGGCGCGGTAAAACTAT	1045
Db	948	AGTTGAAAAAGACTATGCAAAATTATGAGATGAAGCTACTTTCGGTGGCGGAAAAATCAAT	1007
Qy	1046	CCGTGAGGTTATGGTTCAGA-----GCAATAGCCCTGATGAAAAACCCCTAGATT	1096
Db	1008	TCGTGATGGTATGGCTCAAAATCTTAATGTGCAAGAGATGATAAAATGTAGCCGATTT	1067
Qy	1097	AGTCATCACTAACGCGATGATTTCGACTACACCGGATTTTACAAAGCCGACATTCGGAT	1156
Db	1068	AGTTTTAACTAACGCATTAATTATGATTATGACAAGATTGTTAAAGCAGATATCGAAT	1127
Qy	1157	TAAAAACGGCAAAATCCATGGCATTGCAAGCGACGAAACAAGACATGCAAGATGCGGT	1216
Db	1128	TAAAAATGGTTATATTTTAAAGATCGGTAAAGCTGGAAACCCAGATATATATGGATAACGT	1187
Qy	1217	AAGCCCTCATATGGTCTGGTGGGTCAGAAAGCACTAGCAGGGGAAGGTATGATTAT	1276
Db	1188	-----TGACATCATCATTTGGTGCAACAATGATATTTTGTCTGCTGAAGGTAAATTTG	1241
Qy	1277	TACCGTGGGGGAATCGATTACACACCCACTTCTTTTCTCCAACAATTCCTCTACCCG	1336
Db	1242	TACTGCCGCGGTATCGATACACACGCTGCACTTCATCAATCTCTGAACTGAAAGTTGC	1301
Qy	1337	TCTAGCCAAATGGCGTTACAAACCATGTTTGAGCGGCGACAGGTCCTGTAGATGGCAGAA	1396
Db	1302	ACTTGAAGTGGTATTACAACGCATATCGGTGGAGGAACCTGGTCTTCTGAAGGTGCTAA	1361
Qy	1397	TGCACCTACTATCACTCGGCGAAATGGAACCTTGACCGCATGTTGCGCGACAGAGAAGA	1456
Db	1362	AGCAGCTACTGTAACACAGGACCTTGGCATATTCATCGCATGTTAGAACGACGACAGAGA	1421
Qy	1457	GTATTTCTAGAAATGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATGTT	1516
Db	1422	GATCCCTATTAAATGTAGATTACTGGTAAAGGTCAAGCTGTCAATCATACTGCACTTAT	1481
Qy	1517	AGAACAGTAGAAGCGGCGGCGATTGGTTTTTAAATTCGATGAGACTGGGGCACACACC	1576
Db	1482	TGAACAAATTCATGACGCGCTATAGGTCTTAAAGTACATGAAGATTGGGGAGCTACACC	1541
Qy	1577	AAGTGGCATCGATCACTGCTTGACGCTGGCAGATGAATAGATGTGCAAGTTTGTATCCA	1636
Db	1542	TTAGCATTAAGTCAATGCAATAGACGTTGCAGATGAGTTTGATGTTCAAGTCGCTTTACA	1601
Qy	1637	CACCGATACAGTCAATGAGGCGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGCG	1696
Db	1602	TGCAGATACATTAATGAAGCTGGATTATGGAAGATACAAATGGCTCTGTGAAAGATCG	1661
Qy	1697	CGCCATCCATCCCTACACATTTGAGGAGCGGGTGGAGGACATCACCTCATGTTATCAC	1756
Db	1662	TGTTATGCTATGTATCATCTGAAGGAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1721
Qy	1757	CATGGCAGGCGAGCTCAATATTTCTACCCCTCTCTCCACACCCCACTATTCCTCTATACAT	1816
Db	1722	ATCAGCTGCATTTCAAAACATCTTACCTTCTTCTTACAAAACCAACATTAATCTTACCTCA	1781
Qy	1817	TAATACGGTTGCAGAACTTTAGACATGCTCATGATGCGACCACTAGACAAACCGCAT	1876
Db	1782	CAACACTGTAGTGAACAATTTTGAACATGTTTATGATTAATCTCACCATCTTAATGCTTCAAT	1841

RESULT 2

RESOL 2
US-10-793-626-3853

US-10-753-628-3833
: Sequence 3853, Application US/107933626

: Publication No. US20050255478A1

; PUBLICATION NO: 0020
; GENERAL INFORMATION:

APPLICANT: KIMMERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

;; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: Patco

; SEQ ID NO 3853

; LENGTH: 31

TYPE: DNA

ORGANISM

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic

```

OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3853

Query Match 21.2%; Score 612; DB 1; Length 3234;
Best Local Similarity 57.4%; Pred. No. 5.3e-139;
Matches 1211; Conservative 0; Mismatches 845; Indels 5

Qy	543	TTGTTGCGGATAAAGACATCAGAGCTCAATGCGAGCGCAAGAAGTAACCGAACTCTGAGGTGA	602
Db	537	TTGTTAAAAATACTGAAGTAAGAAGTCAATAAACAATCATCCCGAAACGGTGAITGAAGTAA	596
Qy	603	CTAATGAAGGCGCTAAATCCTTGACATGTGGGTAGCCATTTCCACTCTTTTGAAGCTAAACA	662
Db	597	AAAATACAGGCGATAGACCTATACAAGTAGGTTACATTTCCACTTTTTCGAAGCAATA	656
Qy	663	AGGCATTAATAATTCGATTCGTGAATAAGCCTATGGCAAAACGCTAGATATTTCCCTCTGGCA	722
Db	657	AAGCATTTAGAAATTTGATCGTGAGAAAGCATATGGTAACAATTTTGGATATTTCTTGCAGGAG	716
Qy	723	ACACGCTACGCATTTGGGGCAGGACAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGTGGCA	782
Db	717	CTGCAGTGAGATTTGAACCTCGAGATGAATAAAAAAGTACAACTTGTCCGAATATTTCTGGAC	776
Qy	783	GTAAAAAAGTCATTTGGCATGAAACGGGCTTTGTGATTAAC---ATCGCGGATGACGCCATA	839
Db	777	GACGTAAAAATTTATGSGAATTCGGTGGTTTAGTCGATGGCGATATTGACGAAGAAGCGGTAT	836
Qy	840	AACATAAAGCGCTTGACAAGCGCAAAATCTCACGGATTTATCAAGTAAGGAGACTCCCATG	899
Db	837	TCCGTCCAAATGATTCAAATCAAAAACCGCGCGTTAAAAAGATGTCAGGCGAGACAATG	896
Qy	900	-----AAAATGAAAAAACAAGAAATATGTAAA	925
Db	897	CGAATAAAAAAGGTGGTAAATAAGCATGAGTTTTTAAAAATGACACAACTCTCAATACACAAG	956
Qy	926	TACCTACGGACCCACAAAGCGGATAAAGTGCCTTAGGAGATACCGATCTTTGGGCGAGA	985
Db	957	TCTTTATGGACCAACTCTGTAGAGATCTCTGTGAGATTAGGAGATACGAACTTGTTTGACCA	1016
Qy	986	AGTAGAACATGACTATACCACTATGCGGAAGAACTTAAATTTTGGCGCGGTAAACACTAT	1045
Db	1017	AGTTGAAAAAGACTATGCAAAATTTAGAGATGAAGTACTTTTCGGTGGCGGAATAATCAAT	1076
Qy	1046	CCGTGAGGGTATGGGTGAGA-----GCCAATAGCCCTGATGAAACACCCCTAGATTTT	1096
Db	1077	TCGTGATGTTGCTCAAATCCTAATGTGACAAGAGATGATAAAAATGTAGCCGATTT	1136
Qy	1097	AGTCATCACTAAACGGATGATTCGACTACACCGGGATTTACAAAGCCGACATTTGGAT	1156
Db	1137	AGTTTAACTAAACGCATTAATTTATGATATGACAAGATTGTTTAAAGCAGATATCGGAAT	1196
Qy	1157	TAAAAACGGCAAAATCCATGGCATTCGACAGGCGAGGAAAAAAGGACATGCAAGATGCGCT	1216
Db	1197	TAAAAATGGTTATATTTTAAAGTCGGTAAGCTGGAAACCCAGATATAATGGATAACGT	1256
Qy	1217	AAGCCCTCATATGCTCGTGGTGTGGGCACAGAAGCACTAGACGGGAAAGGTATGATTAT	1276
Db	1257	-----TGACATCATATTGGTGCAACAACTGATATTATTTGCTGCTGAAGGTAAAATTTGT	1310
Qy	1277	TACCGCTGGGGGAATCGATTCACACACCCACTTCCTTTCTCCACAACTTCCTCCACCGC	1336
Db	1311	TACTGCCGCGGTATCGATACACACGCTGCACCTCATCAATCCTCGAAACGATGAAGTTGC	1370
Qy	1337	TCTAGCCAAATGGCGTTCAACACCATGTTTGGAGGGCGCACAGGTCTCTAGATGCGACGAA	1396
Db	1371	ACTTGAGATGGTATATCAACGCATATTCGGTGGAGGAACTGGTGCTTCTGAAGGTGCTAA	1430
Qy	1397	TGCGACTACTATCACTCCGGCAAAATGGAACTTGCAACCGCATGTTGCGCGCAGCAGAGA	1456
Db	1431	AGCGACTACTGTAAACACAGGACCTTTGGCATAATTCATCGCATGTTAGAACGACGCGAAGA	1490
Qy	1457	GTATTTCTATGAATGTGGGCTTTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAACCTTGT	1516

Db	1491	GATGCTTATTAACTGTAGGATTTACTGGTAAAGGCTCAAGCTGTCTCAATCATACTGCACTTAT	155
Qy	1517	AGAAACAAGTAGAAGCGGCGCGATTTGGTTTAAATTTGCATGAAGACTGGGGCACAACACC	1576
Db	1551	TGAACAAATTCATGACGGCGCTATAGTCTTAAAGTACATGAAGATTTGGGAGCTACACC	1610
Qy	1577	AAGTGGCAGTCGATCOACTGTCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCA	1636
Db	1611	TTCAGCATTTAAGTCATGCATTTAGACGTTTCAGATAGACTTTGATGTCTCAAGTCGCTTTACA	1670
Qy	1637	CACCGATACAGTCAATCGAGGACAGTTATGTAGATGCACCCCTAAATGCAATGAACGGCGG	1696
Db	1671	TGCAGATACATTAATGAAGCTGGATTTATGGAAGATCAATGGCTCTGTGAAAGATCG	1730
Qy	1697	CGCCATCCATGCCCTACCACATTTAGGAGGCGGGTGGAGGACACTCACCTGATGTTATCAC	1756
Db	1731	TGTATTGCATATGTATCATACTGAAGGAGCTGGTGGTGCATGCACCTGCATTTAATCAA	1790
Qy	1757	CATGGCAGGCGAGTCAATATTTACCCCTCTCCACACCCCCACTATTTCCCTATACCAT	1816
Db	1791	ATCAGCTGCATATTTCAAAACATCTTTACCTCTTCTTACAAACCCCAACATTAACCTTACAC	1850
Qy	1817	TAATACGGTTCAGAAACACTTTAGACATGCTCATGCACATGCCACCACTAGACAAACGCAT	1876
Db	1851	CAACACTGTAGATGAACATTTAGACATGGTTATGATTACTCACCATCTTAATGCTTCAAT	1910
Qy	1877	CGCGAGGATTTTCAAAATTTTCTCAAAGCCGTTATCCGCCCGCGCTCTATCGCGCTGAAGA	1936
Db	1911	ACCAGAAGACATTCGATTTTCAGATTTCTCGTATACGTAAGGAAACTATAGCAGCAGAAGA	1970
Qy	1937	TGTGCTCCATGATATGGGTGTGATTCGCGATGACAAGCTCGGATTCGCAAGCAATGGGGCG	1996
Db	1971	CGTATTACAAGATATGGCGGTATTTAGTATGTTAAGTTCAGATTTCAAAAGCAATGGGACG	2030
Qy	1997	TGCAGGCGAAGTGATTCCTCGAACTTGGCAGACTCGCGATGAAGTAATAAAGAAATTTGG	2056
Db	2031	TGTCGGTGAAGTTGTAAACGCTACTTGGCAAGTTTGCACACGATGAAGAAACCAACGCGG	2090
Qy	2057	TAAGCTCTTCAAGATGCGAAAGATAACGATAATTTCCGCATTAAGCGCTACATCTCAA	2116
Db	2091	ACCATTAGATGGTGACTTTGAAATATCAGATATAATTCGTAATTAACGTTACATTTGCANA	2150
Qy	2117	ATACACTATCAACCCCGCTTTTGACCCACGGCGTGAGCGAGTATATCGGCTCTGTGGAAGA	2176
Db	2151	ATATACAATCAATCCTGCCATTACACATGGTATTTCTGACTATGTTGGATCTGTAGAGAC	2210
Qy	2177	GGGCAAGATCCCGACTTTGGTGTGTGGAATCTCGCCCTTTTGGCGGTAAACCCCAAT	2236
Db	2211	GGGTAAACTTCGCCATTTAGTAAATGTGGGAAACCAAGAAATTCCTCGGTGCCAAACCCGATCT	2270
Qy	2237	CGTGATCAAGGCGGTATGGTGGTCTTCTCTGAAATGGCGATTTCTAACGGCTCTGTGCC	2296
Db	2271	TGTTGTTAAAGGTGGCATGATTAATCAGCAGTAAATGGTGATGCTAATGGCTCCATACC	2330
Qy	2297	CACCTCCCAACCGGCTTTATTACCGCGAAATGTTTTGGGCATCACGGCAAGCGCAATTTGA	2356
Db	2331	AACATCAGAGCTTTTGAAATATCGCAAAATGATATGGTCAAATTTGGTGTAAACATTTACACA	2390
Qy	2357	CACAGCATCATCTTTGTTTCCAAAGTCGCCATATGAATAATGGCGGTGAAGAAAGAGCTGGG	2416
Db	2391	TACTGCTATGACTTTTGTGTTCTTAACACACTGCATATGAAAAACGTAATTTTTCGTCAACTCAA	2450
Qy	2417	CTTAGAGGCCCAAGTCTTACCGGTCAAAAACCTGCGGTAAATCATCACCAAGAAAGACTTCAA	2476
Db	2451	TCTAAAAAGAAATGGTTTCACAGTTTGAATAATTTAGAAATTTAACTAAGGCAGATATGAA	2510
Qy	2477	GTTTCAACGACAAAAACGGCAAAAAATCACCGTCGATCCGAAACCTTCGAGGCTTTTGTAGA	2536
Db	2511	AAATTAATATGCTACACCTAAATATAGATGTAGATCCACAACATATGAGGTATTCGTTGA	2570
Qy	2537	TGGCAAACTCTGCACCTCTAAACCAACCTCGCAAGTGCCTCTAGCCGACGCGCTACACTTT	2596
Db	2571	TGTTAATAAAAAATCAAAAGTGAAGCAGCAACAGAAATTAACCAATTAACAAAGATATCTCTT	2630

QY	2597	CTTCTAGG 2604		QY	1489	GCAATAGCTCTAGCAAAAAACAACCTTGTAGAAACAGTAGAAGCGGGCGCGATTTGTTTTTA	1548
Db	2631	ATTCTAGG 2638		Db	602	GTCAAGCTGTCAATCATCTATGCACTTATTGAAACAAATTCATGACGGCGCTATAGTCTTTA	661
RESULT 3				QY	1549	AATTGCGATGAAGACTGGGCGACACCAACAGTCGATCGATCACTGCTTGAGCGTGGCAG	1608
US-10-793-626-117				Db	662	AAGTACATGAAGATTGGGGAGCTACACCTTCAGCATTTAAGTCATGATTAACGCTTGAATG	721
; Sequence 117, Application US/10793626				QY	1609	ATGAATACGATGTGCAAGTTTGTATCCACCCGATACAGTCAATGAGGACGATTATGTAG	1668
; Publication No. US20050255478A1				Db	722	ATGAGTTTGTATGTTCAAGTCGCTTATCATGCAGATACATTAATGAAGCTGGAATTTATGG	781
; GENERAL INFORMATION:				QY	1669	ATGACACCCCTAAATGCAATGAACGGGCGGCATCCATGCCCTACACATTTGAGGAGCGG	1728
; APPLICANT: KIMMERLY, WILLIAM JOHN				Db	782	AAGATACAATGCGCTGCTGTAAGATCGTGTATTGTCATATGTATCATACTGAAAGGAGCTG	841
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS				QY	1729	GTGAGGAGACACTCACCTGATGTTATCACCATGGCAGGCGAGCTCAATATTTCTACCCCTCT	1788
; FILE REFERENCE: PU3480US				Db	842	GTGGTGGTCATGCACCTGACTTAATCAAAATCAGCTGCATATTCAAAACATCTTACCTTCTT	901
; CURRENT APPLICATION NUMBER: US/10/793,626				QY	1789	CCACCAACCCCACTATTTCCTTATACCATTAATACGGTTGCAGAACACCTTTAGACATGCTCA	1848
; CURRENT FILING DATE: 2004-03-04				Db	902	CTACAAACCCCACTTACCTCACAACTGTAGATGAACATTTAGACATGGTTA	961
; PRIOR APPLICATION NUMBER: 60/164,258				QY	1849	TGACATGCCACCACTAGACAAACGCAATCCGCGAGGATTTTCAATTTTCTCAAAGCCGTA	1908
; PRIOR FILING DATE: 1999-11-09				Db	962	TGATTTACTCACCACTCTTAATGCTTCAATACCAGAAAGACATTTGCATTTGCGATTTCTGTA	1021
; NUMBER OF SEQ ID NOS: 4472				QY	1909	TCGCCCGCGCTCTATCGCGCTGGAAGATGTGCTCCATGATATGGTGTGATCGCGATGA	1968
; SOFTWARE: PatentIn Ver. 2.1				Db	1022	TACGTAAAGGAACTTATAGCAGCAGAAACGCTATTACAAAGATATGGGCGTATTTAGTATGG	1081
; SEQ ID NO 117				QY	1969	CAAGCTCGGATTCGCAAGCAATGGGCGTGCGAGCGGAAGTATTCCTCGAACTTGGCGAGA	2028
; LENGTH: 1716				Db	1082	TAAGTTTCAGATTCACAAGCAATGGGACGTGTCGCTGAACTTTGTACACGCTACTTGGCAAG	1141
; TYPE: DNA				QY	2029	CTCGGTAAGAATAAAAAAGAAATTTGGTAAGCTTCCTGAAAGATGGCAAGATAACGATA	2088
; ORGANISM: Artificial Sequence				Db	1142	TTGCACACCGGTATGAAAGAAACAACGCGGACCCATTTAGATGGTGACTTTTGAATATACGATA	1201
; FEATURE:				QY	2089	ATTTTCGCGATTAAGCGGTACATCTCCAAATACATCTCAACCCCGCTTTGACCCAGCGG	2148
; OTHER INFORMATION: Description of Artificial Sequence: synthetic				Db	1202	ATAATCGTATTAACGTTTACATTTGCAAAATATACAATCAATCCTGCGCATTAACATGTA	1261
; OTHER INFORMATION: nucleic acid sequence				QY	2149	TGAGCGAGTATATCGGCTCTGTGGAAGAGGCAAGATCGCCGACTTGTGTGTGCGAATC	2208
US-10-793-626-117				Db	1262	TTTCTGACTATGTTGGATCTGTAGAAGCGGTAACCTTGCCGATTTAGTAATGTGGGAAC	1321
Query Match 20.1%; Score 578.6; DB 1; Length 1716;				QY	2209	CTGCTTTTTTTGGCGTAAAAACCCAAAATCGTGATCAAGGCGGTATGGTGTCTTCTCTG	2268
Best Local Similarity 59.5%; Pred. No. 4.8e-131;				Db	1322	CAGAAATTTCTCGGTGCCAAACCCCGATCTTGTGTTAAAGGTGGCATGATTAACCTCAGCAG	1381
Matches 1021; Conservative 0; Mismatches 679; Indels 15; Gaps 2;				QY	2269	AAATGGCGGATTTCAACGCGTCTGTGCCACTCCCAACCCGGTTTATTTACCGCGAAATGT	2328
QY	898	TGAAATGAAACCAAGATATGTAATACCTACCGACCCCAACAAAGCGGATAAAGTGC	957	Db	1382	TAAATGGTGATGCTTAATGGGTCCCATACCAACATCAGAGCCTTTGAAATATCGCAAAATGT	1441
Db	8	TTAAATGACAAATCTCAATACACAAGTCTTTATGGACCAACTGAGGAGACTCTGTGA	67	QY	2329	TTGGGCATCAGCGCAAGCGGAAATTTGACACCGCATCACCTTTGTTTCCAAAGTCGCT	2388
QY	958	GCTTAGGAGATACCGATCTTTGGCGCAGAGTAGAATGACTATACCACTATGCCATTTGGCGAAG	1017	Db	1442	ATGGTCAATTTGGTGGTAAACATTACACATATGCTGCTATGACTTTGTTTCTAACACTGCAAT	1501
Db	68	GATTAGGAGATACGAATCTTTTGGCACAAGTTGAAAAGACTATGCAATTTATGGAGATG	127	QY	2389	ATGAAAATGGCGTCAAAAGAAAAGCTGGGCTTAGAGCGCAAGTTCCTACCGGTCAAAAACCT	2448
QY	1018	AACCTAAATTTGGCGGGTAAACATATCCGTGAGGGTATGGGTGACA-----GCA	1068	Db	1502	ATGAAAACGGTATTTATTCGTCACCTCAATCTTAAACGAATGCTTCGACCGATTTAGAATA	1561
Db	128	AAGCTACTTTTCGTGGCGGAAATCAATTCGTGATGGTCTCAAAATCCTTAATGTGA	187	QY	2449	GCCGTAAACATCACCAAGAAAGACTTCAAGTTTCAACGACAAACCGCAAAAAATCACCGTCG	2508
QY	1069	ATAGCCCTCATGAAACACCTAGATTTAGTATCATCACTAACGCGATGATTTATCGACTACA	1128	Db	1562	TTAGAAATTTTAACTAAGCGAGATATGAAAATAATAATATGCTACACCTTAAATATAGATGTAG	1621
Db	188	CAAGAGATGATAAAAAATGTAGCCGATTTAGTTTAACTAACGCAATTAATTTATTTATG	247	QY	2509	ATCCGAAAAACCTTCGAGGTCTTTGTAGATGCAAAACTCTGACACCTCTTAAACCCACTCGC	2568
QY	1129	CGGGATTTTACAAACCGGACATTTGGGATTAATAAACGCAAAATCCATGGCATTTGGCAAG	1188	Db	1622	ATCCACAAACATATGAGGTATTCGTTGATGTAATAAAATCAACAGTGAAGCAGCAACAG	1681
Db	248	ACAGATTTGTTAAGCAGATATCGGAATTAATAATGGTTATATTTTAAAGATCGTAAAG	307	QY	2569	AAGTGCCTCTAGCCCGCGCTACACTTTTCTCTAG 2603	
QY	1189	CAGGAAACAGGACATGCAAGATGGCGGTAAAGCCCTCATATGTCGTGGGTGTGGGCACAG	1248				
Db	308	CTGGAACCCAGATATAATGGATAACGT-----TGACATCATCATTTGGTGCAACACTG	361				
QY	1249	AAGCACTAGCAGGGAAGGTATGATTTATACCGCTGGGGGAATCGATTTACACACCCACT	1308				
Db	362	ATATTAATGCTGCTGAAGGTAAATTTGTTACTGCCGCGGTATCGGATACACAGCTGCAC	421				
QY	1309	TCCTTTCTCCACAAATTCCTACCGCTCTAGCCATGGCGTTTACCACTATGTTGGAG	1368				
Db	422	TCATCAATCTTGAAACAGCTGAAGTTGCACTTGAGAGTGGTATTTACACGCAATTCGGTG	481				
QY	1369	GCGGCACAGTCTCTAGATGGCAGCAATGCGACTACTATCACTCCGGGCAAAATGGAAC	1428				
Db	482	GAGGAATCTGCTGCTTCTGAAGTGTCTTAAGCGACTACTGTAAACACGAGACCTTGGCATA	541				
QY	1429	TGCACCGCATTTGTCGCGCAGCAGAGAGATTTCTATGAATGTGGGCTTTTGGGCAAG	1488				
Db	542	TTCATCGCATGTTAGAAGCAGCAGAGAGATGCTTATTAATGTAGGATTTTACTTGGTAAAG	601				

Db 1682 AATTACCATTAAACACAAAGATACCTTCTATTCTAG 1716

RESULT 4

US-10-485-517-26
; Sequence 26, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James
; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629W0
; CURRENT APPLICATION NUMBER: US/10/485,517
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 1612
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-10-485-517-26

Query Match 15.8%; Score 455; DB 1; Length 1612;
Best Local Similarity 57.5%; Pred. No. 4.1e-101;
Matches 928; Conservative 0; Mismatches 630; Indels 55; Gaps 4;

QY 619 ATCTTGCATGCGGTAGCAATTCACATCTTTTGAAGCTAAACAGGCATTAATAATTGCA 678
Db 6 ACCAATTCAAGTGGGCTCAATTTTCATTTTATGAAGCAAAATGCAGCATTAGATTTCGA 65
QY 679 TCGTGAAGAACGCTATGGCAACCGCTAGATATTTCCCTCTGGCAACACGCTAGCATTTGG 738
Db 66 ACGTGAATGGCATATGGAACAAATTTAGATATTCAGCTGGAGCAGCTTTCGATTGA 125
QY 739 GCAGGACAAACCCGCAAGTCAGATTGATTCCTCTTTGGTGGCAGTAAAGATGATTGG 798
Db 126 ACCTGGGATAAAGAGAGTTCAATTAGTTGAATATGCTGGCAACAGCTAAATTTTGG 185
QY 799 CATGAACGGCTTGTGAA----TAAACATCGCGGATGAACGCCATAACATAAAGCGTTGA 855
Db 186 TTTTCGTGGTATGGTCAAATGGTCTATCGATGATCAGCTGTCTATCGCCCAACTGATGA 245
QY 856 CAAGCGCAATCTCAGCGATTATCA----- 881
Db 246 AATGATGAATATGCGAGTGTATTCGGAGATAACGGTGTGAAGACGTGAATAAAGG 305
QY 882 ---AGTAAAGAGACTCCCATGAATAATGAAAAAACAAGATATGTAATACCTACGGACCC 938
Db 306 AGGAAAAGATCATGAGCTTTAAATATGAACCAAAATCAATATACGAGCTTATACGGTCCA 365
QY 939 ACCAAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGCAGAGTAGAACATGAC 998
Db 366 ACTGTTGGAGATTCATTCGTTTAGGTGATACGAATCTATTTGCTCAATAGAAAAAGAC 425
QY 999 TATACACCTATGGCAGAACCTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATG 1058
Db 426 TATCGCGTTTATGGTGAAGAGCTACTTTTGGTGGTAAATCTATTAGAGACGGTATG 485
QY 1059 GGTCAAGCAATAGCCC-----TGATGAAAAACACCCCTAGATTTTAGTCATCACTAAC 1109
Db 486 GCGCAAAATCCTCGTGTAAACACTGATGACGTGAAGCTTCCAGACCTTGTCTTCTAAT 545
QY 1110 GCGATGATTATCGACTACACCGGGATTTTACAAAGCCGACATTTGGGATTTAAAAACGGCAAA 1169
Db 546 GCGCTTATTATCGATTACGATAAAGTGGTTTAAAGCTGATATAGGCATTTAAAAATGGTTAT 605

RESULT 5

US-10-793-626-4032
; Sequence 4032, Application US/10793626
; Publication No. US20050255478A1

QY 1170 ATCCATGGCATTGGCAAGGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATG 1229
Db 606 ATTTTCGCCATAGGTAATGCGCGCAACCCAGGATAATATGTT-----CGACATT 659
QY 1230 GTCGTGGGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTACCGCTGGGGGA 1289
Db 660 ATTATAGGTTCAACAACAGATATCATTTGCCCTGAAAGTAAATCGTCACTGCTGGTGGT 719
QY 1290 ATCGATTACACACCCCACTTCTTTCTCCACAACAATCCCTACCGCTCTAGCCAAATGCG 1349
Db 720 ATTGATACTCATGTTCAATTTTATTAACTCTGAACAGAGGTCGCATTAGAAAAGTGGT 779
QY 1350 GTTACAACCAATGTTGGAGCGGCACAGTCTCTGTAGATGGCACAATGCGACTACTATC 1409
Db 780 ATTACGACTCATATTTGGTGGTGTACTGCTTCAGAAAGTTCTTAAAGCAACAACCTGTA 839
QY 1410 ACTCCGGGCAATGGAACCTTCACCGCATGTTGCGCGCAGCAGAGAGATTTCTATGAAT 1469
Db 840 ACTCCAGGTCATGGCATAATTCATAGAATGTTAGAAGTGGCGAAGGTTTACCGATTAAT 899
QY 1470 GTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATCTGTAGAACAAAGTAGAA 1529
Db 900 GTCGGTTTACAGTAAAGGACAAGCAACAATCCAATGCACTCATTTGAACAATCAAT 959
QY 1530 GCGGCGCGATTTGGTTTAAATTCATGAAGACTGGGGCAACAACCAAGTGGCATTCGAT 1589
Db 960 GCGGAGCAATTGGATTAAAAAGTACATGAAGACTGGGGTGCAACACCATCTGCTTTGAGT 1019
QY 1590 CACTGCTTGAGCTGGCAGATGAATACGATGTCGAAGTTTGTATCCACACCGATCAGTTC 1649
Db 1020 CATGCATTAGATGTTGCTGATGAATTTGATGTTCAAAATTCATATCATGCAATACTTFA 1079
QY 1650 AATGAGCAGGTTATGTAGATGACACCCCTAAATGCAATGAAGCGGCGCCATCCATGCC 1709
Db 1080 AATGAAGCAGATTATGAAGAACACAATGCTGCTGTTAAAGCCGTGTACTTCATATG 1139
QY 1710 TACCACATTGAGGAGCGGGTGGAGGACACTCACCTGATGTTATCACCATGSCAGCGAG 1769
Db 1140 TACCATACTGAAGTGTGTTGGCGGTGATGCGCTGATTTAAATTAATTCGCTGATTT 1199
QY 1770 CTCAATATTCCTCTCCACACCCCACTATTCCTATATACCATTAACATGCGGTTGCA 1829
Db 1200 TCAAAATATTTACCTTCATCTACAAATCCAACTTTGCCCTTATACACATAATCTGTAG 1259
QY 1830 GAACACTTAGACATGCTCATGACATGCGCACCTAGACAAACGATCCGCGAGGATTTA 1889
Db 1260 GAACATTTAGATGTTAAATGATTACTCACCATTTAAATGCGGCTATTTCCTGAAGATTC 1319
QY 1890 CAATTTTCTCAAGCCGTATCCGCCCGCTCTATCGCGGCTGAAGATGTGTCCTCATGAT 1949
Db 1320 GCATTCCGAGATTCACTGTTTCGTAAGAAACGATTGACAGAGAAGATGTTCTGCAAGAT 1379
QY 1950 ATGGGTGTATCGCGATGACAAAGTCGGATTTCGCAAGCAATGGGCGGTGACGCGAAGTG 2009
Db 1380 ATGGGTGTATTCAGTATGATTAGTTCCGATTCAAAAGCAATGGGCGGTGAGGTGAAGTA 1439
QY 2010 ATTCTCTGAACCTTGGCAGATGCGGATAAGAAATTAAGAAAGATTTGGTAAGCTTCTGAA 2069
Db 1440 ATTAACAGCAATGGCAAGTAGCACATGCGCATGAAAGAACAAACGTTGCTTTAGATGGT 1499
QY 2070 GATTGGCAAGATACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAAC 2129
Db 1500 GATTTTGAACATAATGATAATTAATCGCATCAAAAGTTATATCGCTTAATATATCAATTAAC 1559
QY 2130 CCGCTTTTGACCCACCGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAA 2182
Db 1560 CCAGCAATTACACATGGTATTTCTGAATATGTAGGATCTATCGAGCCGGGCAA 1612

```

; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4032
; LENGTH: 3475
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4032

Query Match          14.7%; Score 424.4; DB 1; Length 3475;
Best Local Similarity 55.1%; Pred. No. 1.6e-93;
Matches 1010; Conservative 1; Mismatches 762; Indels 59; Gaps 7;

QY 7 ATTTCCARCACTCAAGCACATATTGATCCTGTGTGTGGTGGTAAATTCRACCTTGT 66
Db 1653 AATATTTCATCATCAATTACATTTTACTTAAATAATGTTTATGTACGCTGGATATTTT 1712

QY 67 TAAATCTATTATTAATTTTTTAATAATTAATTAATTAATTAATTAATTAATTAAT 126
Db 1713 CATATTATTCAACTTTATACACTTAATTTTAAATAAACATTTGTAATTAGATAAAATAT 1772

QY 127 ATATTAAAAAGTTAATAAAAAAGTAACGAAATAGGACTATAATCCCATTCGCCCTTAAAAAT 186
Db 1773 CAAGTATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1832

QY 187 TTAACAACAAGGA---GTAATAGGTGAACTCACACCAAGAGCAAGAAAAGTTCTTGTT 243
Db 1833 TAATATGAAGAAGGGATAGTTTGCATTTTACACAAAGTGAACAAGACAAATTTGATGAT 1892

QY 244 ATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGAGAGAGGGCTTAAAGCTCAACCAACC 303
Db 1893 AGTTGTAGCTGTGAGGTTGCACGTCGTAGAAAAGCAAGAGGACTTTAAACTTAATCATCC 1952

QY 304 CGAAGCCATTGCTTACATTAGTCCCATATTATGAGCAGAGCGCGCGTGGAAAAAACAAC 363
Db 1953 TGAAGCATTGCTTTAATCAGTGATGAATTAATGAGAGCGCGCGTGATGG---TAAAC 2009

QY 364 CGTTGCCAGCTTATGGAAGAGTGCATGCACTTTTGTAAAAAAGATGAAGTAATGCCCGG 423
Db 2010 GGTAGCTGAACATGAGCTATGGAATAACAATTTTAAACGAGAGAGATGTCATGGATGG 2069

QY 424 GGTGGTAAATATGTTTCCGATCTAGGTGTAGAGCCACCTTTCTCGATGGTACGAAACT 483
Db 2070 CGTAGCTAACTGATTTACAGAACTTGAAATTTGAAGCAACTTTTCCAGATGGTAAAGTT 2129

QY 484 TGTAACTGTGAATTGGCCCATCGAACCCAGATGAGCACTTCAAGCGGGC---GAAGTGAAT 542
Db 2130 AATAACAGTCCATCACCCCAATGTTTAAAGAGGCTATATCAATGATTCCTGGTGAATTA 2189

QY 543 TTGGTTGGCATAAAGACATCGAGCTCAATGTCAGGCAAGAAAGTAACCGAACTTTGAGGTTA 602
Db 2190 TTGTTAAAAATACTGAATAGAGTCAATAAACATCATCCGAAACCGGTGATGAAGTAA 2249

QY 603 CTAATGAAGGCCCTAAATCTTCGATGTGGGTAGCCATTTCCACTTCTTTGAAGCTAAACA 662
Db 2250 AAAATACAGCGATAGACCTATACAAGTAGTGTTCACATTTCCACTTTTTCGAAACAAATA 2309

QY 663 AGGCATTAATTCGATCGTGAAGAACCTATGGCAAAACCCCTAGATATTCCTCTGGCA 722
Db 2310 AAGCATTAGAAATTTGATCGTGAAGAAAGCATATGGTAAACATTTGGATATTTCTGCGAGGAG 2369

QY 723 ACAGCTACGATTTGGGGCAGGACAAACCCGCAAGATGCAAGTTGATTCCTCTTGGTGGCA 782

```

```

Db 2370 CTCAGTGAATTTGAACCTGGAGATGAAAAAAGTACAACTTGTGCAATATTCTTGGAC 2429
QY 783 GTAAAAAAGTATGTCATGAACCGGCTGTGTAAT---AACATCGCGATGAACGCCATA 839
Db 2430 GACGTAAATTTATGATTCGGTGGTGTAGTCGATGGGATATTGAAGAAGAACCGGTAT 2489
QY 840 AACATAAAGCGCTTGACAAGCGGCAAAATCTCACGGATTTATCAAGTAAGGAGACTCCCATG 899
Db 2490 TCCGTCCAAATGATTCAAATCAAAAACGCCCGCTTAAAAACGATGCAGCGGAAGACAATG 2549
QY 900 -----AAAATGAAAAAACAAGATATGTATAA 925
Db 2550 CGAATAAAAAAGGTGTAATAATAGCATGAGCTTTTAAAAATGACAAATCTCAATACACAAG 2609
QY 926 TACCTACGGACCCACCAAGCGGATAAAGTCGCTTAGGAGATACCGATCTTTGGGCGAGA 985
Db 2610 TCTTTATGGACCAACTGTAGGAGACTCTGTGAGATTTAGGAGATACGAACTTGTTTGGACA 2669
QY 986 AGTAGAACATGACTATACCACTTATGCGGAAGAACTTAAATTTGCGCGGGTAAAACTAT 1045
Db 2670 AGTTGAAAAAGACTATGCAAAATTTATGAGATGAAGCTACTTTTCGGTGGCGAAATCAAT 2729
QY 1046 CCGTAGGGTATGGGTGAGA-----GCAATAGCCCTGATGAAAAACACCCCTAGATTT 1096
Db 2730 TCGTGTGTTGCTGCTCAAAATCCTAATGTGCAAGAGATGATAAAAATGTAGCCGATTT 2789
QY 1097 AGTCATCACATAACCGGATGATTATCGACTACACCGGATTTTACAAAGCCGACATTTGGGAT 1156
Db 2790 AGTTTAACTTAACGCATTAATTTATGATTTATGACAAGATGTTTAAAGCAGATATCGGAAT 2849
QY 1157 TAAAAACGGCAAAATTCATGCGCATTTGCAAGCGGAGAAACAAGGACATGCAAGATGGGCT 1216
Db 2850 TAAAAATGGTTATATTTTAAAGATCGGTAAAGCTGGAACCCAGATATAATATGATAACGT 2909
QY 1217 AAGCCTCATATGCTGGGTGGGCAAGAGACATAGCAGGGGAAGGTATGATTAAT 1276
Db 2910 -----TGACATCATCTATGCTGCAACAACATGATTAATTTGCTGCTGAAGSTAAAAATTTGT 2963
QY 1277 TACGCTGGGGGATCGATTTCACACACCCCTCTTCTCCACAACAATTTCCCTACCGC 1336
Db 2964 TACTGCGGGGGTATCGATACACACGTCGACTTTCATCAATCCTGAACAGCTGAAGTTGC 3023
QY 1337 TCTAGCCAAATGGGCTTACAAACCATGTTTGGAGGCGGACAGGCTCCTGTAGATGGCACGAA 1396
Db 3024 ACTTGAGAGTGGTATTACACGCAATATCGGTGGAGGAACCTGCTGCTTCTGAAGGTCTAA 3083
QY 1397 TGGGACTATCATCTCCGGCAAAATGGAATTCGACCGCATGTTGCGCGCAGCAGAGAAGA 1456
Db 3084 AGCGACTACTGTAAACACACGAGCCCTTGGCATATTTCATCGCATGTTAGAAGCAGCAGAGA 3143
QY 1457 GTATTCTATGATGCTGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAACAACACTTGT 1516
Db 3144 GATGCTCTATTAATGTAGGATTTTACTGGTAAAGGCTCAAGCTGTCAATCATCTGCACTTAT 3203
QY 1517 AGAACAGTGAAGCGGCGGATTTGTTTAAATTTGATCAAGACTGGGGCACAACACC 1576
Db 3204 TGAACAAATTCATGCGAGCGCTATAGGTCTTTAAAGTACATGAGATTTGGGAGCTACACC 3263
QY 1577 AAGTGGCATCGATCATCTGCTTGGCGTGGCAGATGAATPACGATGTGCAAGTTTGTATCCA 1636
Db 3264 TTCAGCATTAAGTCATGCAATTAGACGTTGACAGATGAGTTTGTATGTTCAAGTCGCTTTACA 3323
QY 1637 CACCGATACAGTCAATGAGGCGGTTATGTAGATGACACCCCTAAATGCAATGAACGGGG 1696
Db 3324 TGCAGATACATTAATGAAAGCTGGAATTTAATGGAAGATCAATGGCTGCTGTGTAAGATCG 3383
QY 1697 CGCCATCCATCCCTACCACTTGAAGGAGCGGGTGGAGGACACTCACCTGTGTTTATCAC 1756
Db 3384 TGTATTGATGATGATCATCTGAGGAGCTGGTGGTGGTATGATCCCTGATTTAATCAA 3443
QY 1757 CATGGCAGGCGAGCTCAATATTTCTACCCCTCCT 1788
Db 3444 ATCAGCTGCATATTCAAAACATCTTACCTTCTT 3475

```



```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-121

```

```

; Sequence 2; Application US/10834151
; Publication No. US20050244836A1
; GENERAL INFORMATION:
; APPLICANT: TSANG, TAT-KIN
; APPLICANT: MENG, XIANGWEN
; APPLICANT: ZHANG, HONGJUN
; TITLE OF INVENTION: METHODS AND COMPOSITIONS TO DETECT BACTERIA USING MULTIPLEX PCR
; FILE REFERENCE: 37611-95516
; CURRENT APPLICATION NUMBER: US/10/834,151
; CURRENT FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-834-151-2

```

Query Match	7.58;	Score	216.2;	DB	1;	Length	526;
Best Local Similarity	70.9%;	Pred. No.	1.7e-43;				
Matches	287;	Conservative	0;	Mismatches	118;	Indels	0;
Gaps	0;						

QY	900	AAAATGAAAACACAGAATAATGTAAATACCTACGCGACCCACCAAGCGGATAAAGTCGC	959
DB	405	AGATTTAGCAGAAAGANATATGTTCTATGTATGGCCCTACTACAGCGGATAAAGTCAGA	346
QY	960	TTTAGGAGATACCGATCTTTGGCGAGAAGTAGAACATGACTATATACCCTATATGGCGAAGAA	1019
DB	345	TTGGCGGATACAGACTTGATCGCTGGAAGTAGAAACATGACTACACCATTTATATGGCGAAGAG	286
QY	1020	CTTTAAATTTGGCGCGGTAAACATCTCCGTGAGGTATGGTCAGAGCAATATGCCCTGAT	1079
DB	285	CTTTAAATTCGGTGGCGGTAAACCTGAGAGAAGGCATGAGCCAACTCAACACCCCTAGC	226
QY	1080	GAACAACACCCATAGATTTAGTCATCACCTAAACGCGATGATATTCGACTACACCGGGATTTAC	1139
DB	225	AAAGAAGATTGGATCTTAATCATCACCTAACCGCTTTAATCGTGGATTACACCGGTATTTAT	166
QY	1140	AAAGCCGACATTTGGGATTTAAACCGGCAAAATCCATGGCATTTGCCAAGGCAGGAAACAAG	1199
DB	165	AAAGCGGATATTGGTATTTAAAGATGGCAAAATCGCTGGCATTTGGTAAAGCGCGTAAACAA	106
QY	1200	GACATGCAAGATGGCGTAAAGCCCTCATATGGTCGTGGGTGGGCGACAGAAAGCACTAGCA	1259
DB	105	GACATGCAAGATGGCGTTAAACCAATCTTAGCGTAGGTCCTGCTACTGAAGCCTTAGCC	46
QY	1260	GGGGAAGGTATGATTAATTACCGCTGGGGGAATCGATTCAACACC	1304
DB	45	GGTGAAGGTTTGATTCGTAACCTGCTGGTGGTATTGACACACACATC	1

```

RESULT 7
US-10-793-626-121/c
; Sequence 121, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

RESULT 8
US-10-793-626-115
; Sequence 115, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 115
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-115

```

QY 774 TTGTTGGCAGTAAAGTATGGCATCAACGGCGTTGTGAATACATCGCGGATGAA 832
Db 185 ATTCTGGACGAGTAAATTTATGGATTCCGTTGTTAGTCGATCGGATATTGACGAA 243

RESULT 9

US-10-793-626-4078
; Sequence 4078, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4078
; LENGTH: 2259
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4078

Query Match 1.6%; Score 46; DB 1; Length 2259;
Best Local Similarity 48.2%; Pred. No. 0.061;
Matches 158; Conservative 1; Mismatches 166; Indels 3; Gaps 1;
QY 7 ATTTTCCARCACTTCAAGCACATATTGATCCCTGTGTGTGGGTGGTAAATTRCACTTGT 66
Db 1907 AATATCATCATATCAATTAATTTACTAAATAATGTTATGTACGCTGGATATTTT 1966
QY 67 TAATRCATATTAAATTTTTTAATAATTAATTTATATCATATATAATAATTAATTAAT 126
Db 1967 CATATTATTCAACTTTATACACTTAATTTAAAAAACAATGTTAATTAGATAAAATTA 2026
QY 127 ATATTAAAAAGTTAATAAAGTAAGCAATTAAGGACTATAATCCCATTCCTTTAAAT 186
Db 2027 CAAAGTATTATATCTGTATACCTTCATTAATAATAGTTTAAATATAAAATTTGCAAGACAA 2086
QY 187 TTAACACAGGA--GTAATAGTGAACTCACACCAGAGCAAGAAAAGTTCTTGT 243
Db 2087 TAATATGAAGAAGGATAGTTTGCACCTTTACACAACTGTAACAGACAAATTTGATGAT 2146
QY 244 ATATTATGCGGCGAAGTGGCTAGAAAGCGCAAGAGGGCTTAAAGCTCAACCAACC 303
Db 2147 AGTTGTAGCTGTGAGTTGCACGCTGTAGAAAAGCAAGGACTTAACCTTAATCATCC 2206
QY 304 CAAAGCATTGCTTACATTAGTCCCAT 331
Db 2207 TGAAGCACTTGCTTTAATCAGTGATGAT 2234

RESULT 10

US-10-793-626-4389/c
; Sequence 4389, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4389
; LENGTH: 3373
; TYPE: DNA

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4389

Query Match 1.4%; Score 40; DB 1; Length 3373;
Best Local Similarity 56.8%; Pred. No. 2.1;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;
QY 3 GRAGATTTCCARCACTTCAAGCACATATTGATCCTGTGTGGGTGGTAAATTRCAAC 62
Db 2864 GAGGATTATACAGCTTTTGAAAAATATATTAAATTTAAACAGTATTGCTTAATCACAAC 2805
QY 63 TTGTTAATRCATATTAAATTTTTTAATAATTAATTAATTAATTAATTAATTAATTAAT 120
Db 2804 TTGTCATTATAATTTATTTTGTACTATTAAACAATGATTATTGACGTTAATTTA 2747

RESULT 11

US-10-793-626-3926/c
; Sequence 3926, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3926
; LENGTH: 3927
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3926

Query Match 1.4%; Score 40; DB 1; Length 3927;
Best Local Similarity 56.8%; Pred. No. 2.3;
Matches 67; Conservative 3; Mismatches 48; Indels 0; Gaps 0;
QY 3 GRAGATTTCCARCACTTCAAGCACATATTGATCCTGTGTGGGTGGTAAATTRCAAC 62
Db 1836 GAGGATTATACAGCTTTTGAAAAATATATTAAATTTAAACAGTATTGCTTAATCACAAC 1777
QY 63 TTGTTAATRCATATTAAATTTTTTAATAATTAATTAATTAATTAATTAATTAATTAAT 120
Db 1776 TTGTCATTATAATTTATTTTGTACTATTAAACAATGATTATTGACGTTAATTTA 1719

RESULT 12

US-10-793-626-4109/c
; Sequence 4109, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4109
; LENGTH: 3923
; TYPE: DNA

QY 182 AAAATTTAACACGAGTAATAGTGAATCAACCCAAAGAGCAAGAAAGTCTTGT 241
Db 78 ACACCTTTAAGATATAGGAGAAATGAGATGAATCAACCCAAAGAGTATAGTAAAGTTGATG 137
QY 242 TTATATTTAGCGGCGAAGTGTCTGAAGCGCAAGAGAGGCTTAAAGCTCAACCAA 301
Db 138 CTCACACTACGCTGGAGAAATTTGGCTAAAAACGAAAGAAAGGCAATTAAGCTTAACTAT 197
QY 302 CCCGAAGCCATTGCTTACATTAAGTCCCATATATATGGACGAAGCGCGCGTGGAAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTCCCATATATATGGAGAAAGCGAGAGCTGTAAGAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGATGCACTTTTGGAAAAAGATGAAGTAATGCC 421
Db 258 ACTGCGGCTGAATGATGCAAGAAGGCGCACTCTTTTAAAAACCAAGATGATGTGATG 317
QY 422 GGGGTGGGTAAATATGGTTCCCATCTAGGTGTAGAGCCACTTTCCTCATGGTACGAA 481
Db 318 GCGGTGGCAAGCATGATCCATGAAGTGGGTAATGAAGCGATGTTTCTGATGGGACTAAA 377
QY 482 CTTGTAACTGTGAATTTGGCCCATCGAAACGAGATGAGCACTTCAAAAGCGGCGAAGTGAAA 541
Db 378 CTCGTAAACGCTGCAACCCCTATTGAGGCCAATGGTAAATTAAGTTCCTGGTGAAGTTG 434
QY 542 TTTGGTTGGATPAAAGACATCGAGCTCAATGAGGCAAGAAAGTAACCGAACTTTGAGGTT 601
Db 435 TTTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGAAGTT 494
QY 602 ACTAATGAAGGCGCTAAATCCTTGATGTGGGTAGCCACTTCCACTTCTTGAAGCTAAC 661
Db 495 AAAAATGTTGGCGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTGAAGTGAAT 554
QY 662 AAGGCACATAAAATTCGATCGTGAAGAAAGCCTATGGCAAGCGCTAGATATTCCTCTGGC 721
Db 555 AGATGCGCTAGACTTTGACAGAGAAAAAACTTTTCGGTAAACGCTTAGACATTGGACGGG 614
QY 722 AACACGCTACGATTGGGCGAGACAAACCCCGAAAGTCAGTTGATTCCTCTTGTGGC 781
Db 615 ACAGCGTAAAGATTGAGCTGCGGAAGAAATAATCGTAGAATTGATTAACATTGGCGGT 674
QY 782 AGTAAAAAGTGAATGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAA 841
Db 675 AACAGAAGAATCTTTGGATTAAACGCAATTTGGTTGATAGACAAGACACAACGAAACAA 734
QY 842 CATAAAGCGCTTGACAAGCGAAATCTCACGGATTT- 877
Db 735 AAAATTTGCTTTACACAGACTAAAGAGCGTGGTTTTCATGGCGCTAAAAGCGATGACAAC 794
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA--ATGAAAAAACAAGATAATGTAAAT 926
Db 795 TATGTAATAAACAAATTAAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTCT 854
QY 927 ACCTACGACCCACAAAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGGTCTCTACTACAGCGGATAAAGTGAATGGGCGATACAGACTTGATCGCTGAA 914
QY 987 GTAGAACATGACTATACCACTATGGGAGACTTAAATTTGGCGGGTAAACATATC 1046
Db 915 GTAGAACATGACTATACCACTTATGGCAAGAGCTTAAATTTGGTGGCGGTAAACCCCTA 974
QY 1047 CGTGGGGTATGGGTACAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCATCACT 1106
Db 975 AGAAGAGGATGAGCCAAATCTAACACCCCTAGCAAGAGAGTTGATTTAATATCACT 1034
QY 1107 AACCGGATGATTATCGACTACACCGGGAATTAACAAAGCCGACATGGGATTAATAAAGCGG 1166
Db 1035 AACGCTTTAATCGTGAATTACACCGGTATTTTAAAGCGGATTTGTTAATAAGATGGC 1094
QY 1167 AAAATCCATGGCATTTGGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCAT 1226
Db 1095 AAAATCGCTGGCATTTGGTAAGCGGTAACAAAGACATGCAAGATGGCGTTAAAAACAAT 1154
QY 1227 ATGGTCTGGTGGGCACAGAAGCACTAGCAGGGGAAGGTATGATTATTATCCGCTGGG 1286

Db 1155 CTTAGCGTAGGTCTCTACTAGAGCCCTTAGCCGGTGAAGTTTGATCGTAACGGCTGGT 1214
QY 1287 GGAATCGATTACACACCCACTTCTCTTCTCCACAACTTCCCTACCGCTCTAGCCCAAT 1346
Db 1215 GGTATTGACACACATCCACTTCATTTTACCCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGTTTACAAACATGTTTGGAGCGGCAAGGTCTCTGTAGATGGCAGAAATGCGACTACT 1406
Db 1275 GGTGTAAACCAATGATGGTGGAAACCGGCTCTGCTGATGGCACTTAATGCGACTACT 1334
QY 1407 ATCACTCCGGGCAATGGAATTTGCACCCGATGTTGCGCCAGCAGAGAGATATCTATG 1466
Db 1335 ATCACTCCAGGCAAGAAATTTAAATGATGCTCAGAGCGGCTGAAGAAATTTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAAACAATTTGTAGAACAAAGTA 1526
Db 1395 AATTTAGTTCCTTGGCTAAAGGTAGCGTTCTTAACGATGGGAGCTTAGCCGATCAATTT 1454
QY 1527 GAAGCGGCGCGATTGGTTTAAATTTGCATGAAGACTGGGGCACAACCAAGTGCAGTC 1586
Db 1455 GAAGCGGTGCGATTGGCTTTAAAAATTCAGAAAGACTGGGGCACACTCTCTTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGCAAGTTTGTATCCACACCGATACA 1646
Db 1515 AATCATGCGTTAGATGTTGCGGCAAAATACGATGTCAAGTCTGCTATCCACACAGACACT 1574
QY 1647 GTCAAATGAGGACAGTTATGTAGATGACACCTTAATGCAATGAACGGGCGCGCATCCAT 1706
Db 1575 TTGATGAAGCCGGTTGTGTAGAGACACTATGGCTGCTATTGCTGGACGACTATGGAC 1634
QY 1707 GCCTACCACTATGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGC 1766
Db 1635 ACITTTCCACACTGAAGCGCTGGCGGCGACACGCTCCTGATATTTAAAGTAGTCGGT 1694
QY 1767 GAGCTCAATATTTACCTCTCCACCCAGCTATTTCCCTATACCATTAATACGGTT 1826
Db 1695 GAACACAACTTTCTCCGCTTCCACTAACCCCAACCATCTCTTCCACGTGAATACAGAA 1754
QY 1827 GCAGAACCTTTAGACATGCTCATGACATGCCACCTAGACAAACGATCGCGAGGAT 1886
Db 1755 GCAGAGACATGACATGCTTATGTTGTGCCACCACTTGGATTAAGCAATTAAGAAAGAT 1814
QY 1887 TTACAATTTTCTCAAGCGCTATCCGCGCGGCTCTATCCGCGCTGAAGATGTGCTCCAT 1946
Db 1815 GTTCAGTTCCGCTGATTCAGGATCCGCTCAACCATTTGCGGCTGAAGACACTTTGCTAT 1874
QY 1947 GATATGGGTGTATTCGCGATGACAAAGCTCGGATTCGCAAGCAATGGGCGGTGAGCGGAA 2006
Db 1875 GACATGGGGATTTTCTCAATCACCACTCTGACTCTCAAGCGATGGGCGGTGCGGTGAA 1934
QY 2007 GTGATTCCTCGAATTTGGCAGACTGCGGATAAGATTAAGAAAGATTTGGTAAGCTTCCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAAAGAAAGATTTTGGCGCTTGAAA 1994
QY 2067 GAAGATGGCAAGATAACGATAATTTCCGATTAAGCGCTACATCTCCAAATACACTATC 2126
Db 1995 GAAGAAAAAGCGAATAACGAACTTCAAGATCAACCGTACTTCTTAATACACCATTT 2054
QY 2127 AACCCCGCTTTGACCCCAAGCGGTGAGGAGTATATCGGCTCTGTGGAAAGAGGCGAAGATC 2186
Db 2055 AACCCAGCGATCGCTCATGGGATTTAGCGAGTATGTAGGTTCAGTAGAAAGTGGGCAAGTG 2114
QY 2187 GCGCACTTGTGTGTGGAAATCTCGCTTTTGGCGTAAAAACCCAAATCGTGTATCAA 2246
Db 2115 GCTGACTTGTATTGTGGAGTCCAGCATTTCTTTGGCGTGAACCCCAACATGATCATCAAA 2174
QY 2247 GCGGTATGTGTCTTCTCTGAAATGGCGATTTCAACGCTCTGTGCCCACTCCCAA 2306
Db 2175 GCGGATTCATTGGTTAAGCCAAATGGGCGATGCGAAGCGTTCTATCCCTACCCACAA 2234
QY 2307 CCGGTTTATACCGCGAAATGTTTGGGCGATCACGCGAAGGCGGAAATTTGACACCGACTC 2366

Db 2235 CCGGTTTATTACAGAGAAATGTTGCTCATCATGTTAAAGCTAAATACGATGCAAAACATC 2294
QY 2367 ACTTTTGTTCCTCAAGTCGCTATGAAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGC 2426
Db 2295 ATTTTGTGTCTCAAGGGCTTATGACAAAGGCATTAAAGAGATTTAGGACTTGAAGA 2354
QY 2427 CAAGTTCTACCGGTCAAAAACCTCGGTAACATCAACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAGTGTGCGGTAAAAAATGAGAAATATCACTAAAAAAGACATGCAATTCACGAC 2414
QY 2487 AAACGGCAAAATCACCGTCGATCCGAAAACCTTCAGGTCCTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACAATGAAGTCAATCTCTGAAAACCTTACCATGTTGTCGTGATGGCAAGAA 2474
QY 2547 TGCACTCTAAACCACTCGCAAGTCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
Db 2475 GTAACTTCTAAACAGCAATAAAGTAGCTTGGCGCAACTCTTTAGCAATTTTCTAGG 2532

RESULT 2

PCT-US96-05800-1
; Sequence 1, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: OraVax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2735 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-05800-1

Query Match 37.3%; Score 1075.2; DB 6; Length 2735;
Best Local Similarity 66.0%; Pred. No. 5.8e-267;
Matches 1622; Conservative 0; Mismatches 798; Indels 38; Gaps 3;
QY 182 AAAATTTAACACAGGAGTAAATAGTGAAACTACACCCCAAGAGCAAGAAAAGTTCTTG 241
Db 78 ACTTTTAGAATAGGAGATGAGATGAACTACCCCAAAAGAGATTAGATGATTTGATG 137

QY 242 TTATATTATGCGGGCGAAAGTGCTAGAAAGCCCAAGACGAGGGCTTAAAGCTCAACCAA 301
Db 138 CTCACACTACGCTGGAGAATTGGCTAAAAAAGCAAGAAAAGGCAATTAAGCTTAACTAT 197
QY 302 CCCAGAGCAATGCTTACATTAGTGCCTATATATGACGAAGCGCGCTGCAAAAAA 361
Db 198 GTAGAAGCAGTAGCTTTGATTAGTGCCTATATATGGAAGAACGAGAGAGCTGTAAAAAG 257
QY 362 ACCGTTGCCAGCTTATGGAAGAGTGCATGCTTTTGTGAAAAAAGATGAAGTAATGCCCC 421
Db 258 ACTCGGCTGAATGATGCAAGAGGGCGCACTCTTTTAAACCAGATGATGTATGGAT 317
QY 422 GGGGTGGGTAATATAGTTCCCGATCTAGGTGTAGAACCCACTTTCCTGATGTACGAAA 481
Db 318 GCGGTGGCAAGCATGATCCATGAAGTGGGTAATGAAGCGATGTTCTCTGATGGCACTAA 377
QY 482 CTTGTAATCTGAATTTGGCCCATCGAACCAAGATGAGCACTTCAAGCGGGCGAGTGAAA 541
Db 378 CTGCTAACCGTGCATACCCCTATTGAGGCCAATGGTAAATTTAGTTCTCTGTTGAGTTG -- 434
QY 542 TTTGTTGCGATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAACCGCACTTGAGGTT 601
Db 435 TTCTTAAAAAATGAAGACATCACTATCAACGAAGGCAAAAAAGCCGTTAGCGTGAAAGTT 494
QY 602 ACTAATGAAGGCGCTTAAATCTTTCATGTTGGGTAGCCATTTCCACTTCTTTGAAGCTAAC 661
Db 495 AAAAATGTTGGCGACAGACCGGTTCAAATCGGCTCACACTTCCATTTCTTTGAAGTGAAT 554
QY 662 AAGGCATAAAAATTCGATCGTGAAGAAAGCTTATGCAAAACGCGCTAGATATTCCTCTGGC 721
Db 555 AGATGCTTAGACTTTGACAGAGAAAAAACTTTTCGGTAAACGCTTAGACATTCGAGCGGG 614
QY 722 AACAGCTACGATTTGGGCGAGGACAAACCCCAAGTGCAGTTGATTTCTCTTGGTGGC 781
Db 615 ACAGCGTAAAGATTTGAGCTTGGCGAAGAAAAATCCGTAGAAATGATGACATTTGGCGGT 674
QY 782 AGTAAAAAAGTATTGGCATGAACCGGCTTGTGAATAACATCGCGATGAACGCCATAAA 841
Db 675 AACGAGAATCTTTGGATTTAACCCATTTGTTGATAGCAAGCAGACAAACGNAAGCNA 734
QY 842 CATAAAGCGCTTGACAAAGCGGAAATCTCACCGATTT----- 877
Db 735 AAAATGCTTTACACAGAGCTAAAGAGCGTGTGTTTTCATGGCGCTAAAAAGCGATGACAAC 794
QY 878 -----ATCAAGTAAGGAGACTCCCATGAAA-----ATGAAAAACAAAGATATGTAAT 926
Db 795 TATGTAAAAACAATTAAGGAGTAAGAAATGAAAAAGATTAGCAGAAAAAGATATGTTTCT 854
QY 927 ACCTACGACCCCAAGAGCGATAAAGTGGCTTGGAGATACCGATCTTTGGGCGAGAA 986
Db 855 ATGTATGTCCTACTACAGGCGATAAAGTGAATTTGGGCGATACAGATTTGATCGCTGAA 914
QY 987 GTAGAACATGACTATACCACCTTATCGCGAAGAACTTAAATTTGGCGGGGTAAAACTATC 1046
Db 915 GTAGAACATGACTACACCATTTATGCGGAGAGCTTAAATTCGTGGCGGTAAAAACCTTA 974
QY 1047 CGTAGGGTATGGGTACAGACAAATAGCCCTGATGAAAAACACCCCTAGATTTAGTCATCT 1106
Db 975 AGAGAAGGCATGAGCCAACTTAAACACCTTAGCAAAAGAGATTGGATTTAATATCACT 1034
QY 1107 AACGGGATGATTGACTACACCGGATTTTCAAGCGGATGACATGGGATTTAAAAACGGC 1166
Db 1035 AACGCTTTTAAATCGTGATTTACACCGGATTTTATAAGCGGATTTGGTATTTAAAGATGGC 1094
QY 1167 AAAATCCATGCTATGGCAAGCAGCAAAACAGGACATGCAAGATGGCGCTTAAGCCCTCAT 1226
Db 1095 AAAATCGTGGCATTTGGTAAAGCGGTAACAAAGACATGCAAGATGGCGCTTAAAAACAAT 1154
QY 1227 ATGTTGCGGGTGGGCAAGAACACTAGCAGGGGAGGTATGATTAATTTACCGCTGGG 1286
Db 1155 CTTAGCGTAGTCTCTGCTACTGNAAGCCCTTAGCCGCTGAAGGTTTGATCTGTAAGCGCTGGT 1214

QY 1287 GGAATCGATTACACACCCACCTCTCTTCTCCACAAACAATTCCTACCGCTCTAGCCAAT 1346
Db 1215 GGTATTGACACACATCCACTTCAATTTACCCCAACAAATCCCTACAGCTTTTGCAGC 1274
QY 1347 GCGGTTACAACTATTTGGAGCGGCACAGGTCCTGTAGATGGCAGCAATCGGACTACT 1406
Db 1275 GGTGTAAACAACCATGATTGGTGGTGAACCGGTCCTGTGATGGCACTAATGGCACTACT 1334
QY 1407 ATCACTCCGGGCAAAATGGAACCTTGCACCGCATGTTCGGCGCAGCAGAGAATTTCTATG 1466
Db 1335 ATCACTCCAGGCAGAGAAATTTAAATCGATGCTCAGAGCGCTGAAGAAATTTCTATG 1394
QY 1467 AATGTGGGCTTTTGGGCAAGCAATAGCTCTAGCAAAAACAATTTGTAGAACAAATTA 1526
Db 1395 AATTTAGGTTCTTGGCTAAAGGTAACGCTTTTAAACGATGCGAGCTTAGCCGATCAAT 1454
QY 1527 GAAGCGGCGGATTTGGTTTAAATTTGATGATGAGACTGGGACACACCAAGTGGATC 1586
Db 1455 GAAGCGGCTGCAATTTGGCTTTAAATTTACGAAGACTGGGGCACCACTCTTCTGCAATC 1514
QY 1587 GATCACTGCTTGAGCGTGGCAGATGAATACGATGTCAAGTTTGTATCCACACCGATACA 1646
Db 1515 ATCATGCGTTAGATGTTGGGACAAATAGATGTCAAGTGCCTATCCACACAGACT 1574
QY 1647 GTCAATGAGCGGTTATAGATGACACCTTAAATGCAATGAACGGGCGCGCCATCCAT 1706
Db 1575 TTGAATGAAGCGGTTGTGTAGAAGACACTATGGCTGCTATTGCTGGAGCACTATGCAC 1634
QY 1707 GCTACCAATTTGAGGAGCGGTGAGGACACTCACTGATGTTATCACCATGGCAGC 1766
Db 1635 ACTTTCCACACTGAAGCGCTGGCGGCGACACGCTCTGATATTATTAAGTAGCCGCT 1694
QY 1767 GAGCTCAATATTCTACCTCTCCACACCCCACTATTCCCTATACCATTATACGGTT 1826
Db 1695 GAACACAACTTTCTCCGCTTCCACTAACCCACCACTCCCTTACCGTGAATACAGAA 1754
QY 1827 GCAGAACACTTAGACATGCTATGATCGCACCACTGACAAACGCAATCCGCGAGGAT 1886
Db 1755 GCAGAGCACTGACATGCTTATGGTGTGCCACCACTTGATGATAAAGCAATTAAGAAGAT 1814
QY 1887 TTCAATTTTCAAAAGCGTATCCGCGCGGCTCTATCGGCTGAAAGTGTCTCAT 1946
Db 1815 GTTCAGTTGCTGATTTCAAGGATCCGCTCAAAACCAATTTGCGCTGAAGACACTTTGCAT 1874
QY 1947 GATATGGGTGATCGCGATGACAGCTCGAATTCGACGAATGGGCGTGCAGCGAA 2006
Db 1875 GACATGGGGAATTTCTCAATCACTGATCTCAAGCGATGGCGCTGFGGTGAA 1934
QY 2007 GTGATTCCTCGAATTTGGCAGACTGCGGATGAAGATAAAGAAATTTGGTAAAGCTTCT 2066
Db 1935 GTTATCACTAGAACTTTGGCAACAGCTGACAAAAACAGAAAGAAATTTGGCGCTTGAA 1994
QY 2067 GAAGATGGCAAGATAACGATAATTTCCGATTAAGCGCTTACATCTCCAAATACATATC 2126
Db 1995 GAAGAAAAAGCGATACACCACTTCAAGATCAACCGCTTCTGTCTAAATACCACT 2054
QY 2127 AACCCGCTTTGACCCACCGCTGAGCGATATCGGCTCTGTGGAAGAGGCGCAATC 2186
Db 2055 AACCCAGCGATCGCTCATGGATTAGCGAGTATGTAGGTTCACTAGAAAGTGGGCAAGTG 2114
QY 2187 GCGACTTGGTGTGGGAATCTGCTCTTTTGGCGTAAACCCCAAAATCGTGATCAAA 2246
Db 2115 GCTGACTTGGTATTTGGAGTCAGCACTTTTGGCGTGAACCCCAACATGATCATCAA 2174
QY 2247 GCGGATGCGTGTCTCTCTGAAATGGGCGATTTCTAACCGCTGTGTGCCCACTCCCCAA 2306
Db 2175 GCGGATTCATTCGTTAAGCCAAATGGGCGATGCGAAGCTTCTATCTCCCTACCCACAA 2234
QY 2307 CCGGTTTATACCGCAATATTTTGGCATCAGCGCAGCGGAATTTGACACAGATC 2366
Db 2235 CCGGTTTATACAGAGAAATGTTTCGCTCATCATGGTAAAGCTAAATACGATGCAAAATC 2294
QY 2367 ACTTTGTTTCCAAAGTCGCTATGAAATGGCGTGAAGAAAGCTGGGCTTTAGAGCGC 2426

Db 2295 ACTTTTGTGTCTCAAGCGCTTATGACAAAGGCAATTAAGAGAAATTAGGACTTTGAAAGA 2354
QY 2427 CAAAGTTTCTACCGGTCAAAAACCTCCGTAAACATCACCAAGAAAGACTTCAAGTTCAACGAC 2486
Db 2355 CAAAGTTTGGCGGTAAAAAATTCAGAGAAATATCACTAAAAAGACATGCAATTTCAACGAC 2414
QY 2487 AAAACGGCAAAATCAACCGTCGATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTC 2546
Db 2415 ACTACTGCTCACATTTGAAGTCAATCTCTGAAACTTACCATGTGTTTCGTGGATGGCAAGAA 2474
QY 2547 TGCACTCTAAACCCACCTCGCAAGTGCCTCTAGCCAGCGCTACACTTTTCTTAGG 2604
Db 2475 GTAACTTTAAACAGCCCAATAAAGTAGAGCTTGGCGCAACTCTTTAGCAATTTTCTAGG 2532

RESULT 3

US-08-467-822-19
; Sequence 19, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno

OTHER INFORMATION: sequence."									
FEATURE:									
NAME/KEY: misc feature									
LOCATION: 756..759									
OTHER INFORMATION: /standard_name= "Shine-Dalgarno									
OTHER INFORMATION: sequence."									
US-08-467-822-19									
Query Match 36.9%; Score 1062.6; DB 2; Length 2619;									
Best Local Similarity 65.5%; Pred. No.1e-263;									
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;									
QY	192	ACAAGGAGTAAATAGTGAAGTCAACCCAAAGAGCAAGAAAGTCTCTTGTATATATG	251						
DB	29	ATAAGGAGTTAGGATGAATCAACGCTAAAGAACTAGACAAGTTAATGCTCCATTATG	88						
QY	252	CGGGCGAAGTGGCTAGAAAGGCGAAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCA	311						
DB	89	CGGCGAGATTGGCGAAGAAACGCTTGGCGGTGGTGAATCAATTTACACCGAAGCGG	148						
QY	312	TTGCTTACATTTAGTGCCCATATTTATGACGAAGCGCGCTGGGAAAAAAACCCTTGCC	371						
DB	149	TCGGCTCATTAGCGGCGGTGTGATGGAAGAGCGCGTGATGGTAATAAAGCGTGGCG	208						
QY	372	AGCTTATGGAAGGTGCATGCACCTTTTGAATAAAGATGAAGTAAATGCCGGGTGGGTA	431						
DB	209	ATTGTATGCAAGAGGCGAGGACTTGGCTTAAAAAAGAAATGTGTGACGCGCTAGCA	268						
QY	432	ATATGTTCCCGATCTAGTGTGTAAGCGACCTTTTCTGTGATGGTACGAACTTTGTA	491						
DB	269	GCATGATTTCATGAAGTGGGATTTGAAGCTAACTTCCCGATGGAACCAAGCTTGTAA	328						
QY	492	TGAATGGCCCATCGAACAGATGAGCAGTCAAGCGGCGGAAGTGAATTTGGTTGCG	551						
DB	329	TCACACTCCGGTAGAGGATAATGGCAAAATAGCCCCCGCGGAGTCTTCTTAA--AA	385						
QY	552	ATAAAGACATCGAGCTCAATGCGAGCAAGAAAGTAAACCGAACTTGAGGTTACTAAT	611						
DB	386	ATAGGACATTACTATTAAACCGCGCAAGAGCCATTAGCTTGAAGTGAATAAAG	445						
QY	612	GGCTTAAATCTTGATGTGGGTAGCCATTTCACATTTCTTTGAAGCTAACAAAGCACT	671						
DB	446	CGCATCGTCTGTGCGAGTGGATCAATTTCCACTTCTTCCGAAGTGAATTAAGCTTGG	505						
QY	672	AATTCGATCGTGAAGAGCCTATGGCAAGCCCTAGATATTCCTCTGGCAACAGCTAC	731						
DB	506	ACTTCGATCGCAAAAGCTTTTGCAAGCCCTAGACATTTGCATCTTGGACAGCGGTGC	565						
QY	732	GCAATGGGCGAGGACAAACCGCAAGTGCAGTTGTATTCCTCTTGGTGGCAGTAAAAA	791						
DB	566	GCTTTGAACCGGGAGGAAAAAGTGTGAACCTCATTTGACATCGGCGGGAAATAGCGCA	625						
QY	792	TGATTGGCATGAACGGGCTTTGTGAATAACATCGCGGATGAACGCGCATAAAAAT	851						
DB	626	TCATAGCTTTAATCTTTGTGTGATCGCAAGCCGATGCCGATGGTAAAAAACTCGSCT	685						
QY	852	TTGACAAAGGCGAAATCTCAGGATTT-----	877						
DB	686	TAATAACGCTTAAAGAAAAAGTTTGGGTCTGTAAACTGCGGTTGTGAAGCGCACTAA	745						
QY	878	----ATCAAGTAGGAGACTCCCATGAAATGAAATAACAGAAATATGTAAATACCTACG	933						
DB	746	ATAAACAAATAGGAAAAACCATGAAAAAGATTTTCCGAAAAAGAAATATGTTCTATG	805						
QY	934	GACCCCAAGGCGATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAAGTAGAAC	993						
DB	806	GTCCCACTACGGGGATCGTTAGACTCGGCGACACTGATTGATCTTAGAAGTGGAGC	865						
QY	994	ATGACTATACCACTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCCGTGAGG	1053						
DB	866	ATGATTGCACCACTTATGGTGAAGAGATCAAAATTTGGGGGGGTAAAACTATCCGTG	925						
QY	1054	GTATGGGTAGAGCAATAGCCCTGATGAAAAACCCCTAGATTTAGTCTACTTAACGGCA	1113						

DB	926	GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGATTGGTGTCTCACTAAACCCC	985						
QY	1114	TGATTATCGACTACACCGGGATTTTACAAAGCCGCACATTGGGATTTAAAAACGGCAAAATCC	1173						
DB	986	TCATTGTGGACTATAGGGCAITTTACAAAGCCGACATTTGGATTTAAAGACGGCAAGATTG	1045						
QY	1174	ATGGCAATTGGCAAGGCGAGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGTGTCG	1233						
DB	1046	CAGGCATTGGCAAGGCGAGCAATAAGGACATGCAAGATGGCGTAGATAATAATCTTTTGGC	1105						
QY	1234	TGGGTGTGGGCACAGAACGACATAGCAGGGGAAGGTATGATTATTACCGCTGGGGGAATCG	1293						
DB	1106	TAGTCTCTGCTACAGAGGCTTTGGCAGCTTGAGGGCTTGATTGTAACCGCTGGTGGCATCG	1165						
QY	1294	ATTACACACACCCACTTTCTTCTCCACAACAAATTCCTTACCCTCTTAGCCCAATGGCGGTTA	1353						
DB	1166	ATACGCATATTTCATCTTATCTCTCCCAACAATCCCTACTGCTTTTTCAGCGGGGTTA	1225						
QY	1354	CAACCATGTTTGGAGCGGCACAGGTCTCTGTATAGTGGCAGCAATGCGACTTACTCATCTC	1413						
DB	1226	CAACCATGATTGGAGGAGSCACAGGACCTGCGGATGGCAGCAATGCGACCACCATCACTC	1285						
QY	1414	CGGCAAAATGGAACCTTGCAACCGCATGTTGCGCGCAGCAGAGATTTCTATGAATGTGG	1473						
DB	1286	CGGACGCGCTAATCTTAAAAAGTATGTTGCGTGCAGCCGAAAGAAATACGCCATGAATCTAG	1345						
QY	1474	GCTTTTTCGGCAAGGCAATAGCTCTTAGCAAAAAACAACCTTGTAGAACAAAGTAGAAGCGG	1533						
DB	1346	GCTTTTTCGGCTAAGGGGAATGTTCTTACGAACCCCTTTTACGCGATCAGATTGAAGCAG	1405						
QY	1534	GCGCATTTGGTTTAAATTTGCATGAAGACTGGGGCAACAACCAACATGGCGATCGATCACT	1593						
DB	1406	GGGCGATTTGGTTTAAATTCACGAGACTGGGGAAGCACACTGCGAGCTATTTCACCACT	1465						
QY	1594	GCTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG	1653						
DB	1466	GCCTCAATGTGCGCCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCTTAAACG	1525						
QY	1654	AGGCGAGTTATGTAGATGACACCTTAAATGCAATGAACGGGCGCGCATCCATGCCTACC	1713						
DB	1526	AGGCGGCTGTGTAGAAGACACCTTAGAGCGGATTTGCCGGGCGCACCTCCATACCTTCC	1585						
QY	1714	ACATTGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATGGCAGGCGAGCTCA	1773						
DB	1586	ACACTGAAGGGCTGGGGTGGACAGCTCCAGATGTTATCAAAATGGCAGGGAATTTA	1645						
QY	1774	ATATTCTACCTCTCCACACCCCACTATTTCCTTATACCATTAATACGGTTGCGAAGC	1833						
DB	1646	ACATTCTACCCGCTCTACTAACCCGACCATTTCTTTACCAAAAAACACTGAAGCCGAGC	1705						
QY	1834	ACTTAGACATGTCTATGACATGCCACCTAGACAAACGATCCCGAGGAGTTTACAAT	1893						
DB	1706	ACATGGACATGTTAATTTGGTGTGCCACCACTTGGATAAAAAAGTATCAAGGAAGATGTGCAGT	1765						
QY	1894	TTTCTCAAAGCGGTATCCGCCCGGCTCTATCGCGCTGAAGATGTCTCCATGATATGG	1953						
DB	1766	TGCGCATTCGAGGATTCGCCCCCAAACTATCGGGCTGAAGACCAACTCCATGACATGG	1825						
QY	1954	GTGTGATCGCGATGACAAAGCTCGGATTTCGAAGCAATGGGGGTGCGAGCGAAGTGAATTC	2013						
DB	1826	GGATCTTTTCTATCACCGACTCCGACTCTCAGGCTATGGGACGCGTAGGCGAGGTGATCA	1885						
QY	2014	CTCGAACTTGGCAGACTCGGATAGAAATAAAAAAGAAATTTGGTAAGCTTCTTGAAGATG	2073						
DB	1886	CAGGCATTTGGCAGACAGCAGACAAAAACAAAAGAGTTTGGGCGCTTGAAGAGAGAAA	1945						
QY	2074	GCAAGATAACGATTAATTTCCGCATTAGCGCTACATCTCCAAATACACTATCAACCCCG	2133						
DB	1946	AAGCGATAACGACAACTTCGCGATCAAAACGCTACATCTCTAAATACACCATCAACCCCG	2005						
QY	2134	CTTTGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGTGAAGAGGGCAAGATCGCCGACT	2193						

Db 2066 GGATCGCGCATGGGATTTCTGACTATGTGGGCTCTGTGGAAGTGGGCAATATACCGCGACC 2065
Qy 2194 TGGTGTGTGGGAATCTCGCTTTTGGGGTAAACCCCAAAATCGTGATCAAAAGCGGTA 2253
Db 2066 TGGTGTGTGGAGTCCGGCTTCTTGGGCAATTAAGCCCAATATGATTTAAAGGGCGGAT 2125
Qy 2254 TGGTGTGTCTCTGAAATGGGCGATTTCAACCGGCTGTGCGCACTTCCCAACCGGTTT 2313
Db 2126 TTATTGCGCTCTCTCAATGGGCGATGCGCAATGCGTCTATTCCCAACCGCTCAGCCGCTCT 2185
Qy 2314 ATTACCGCGAATGTTTGGGCGATCGGCGAAGCGGAAATTTGACACGAGATCACTTTG 2373
Db 2186 ATTACCGTGAATGTTTGGGACCATGCGGGAACCAATTCGACACCAATATCACTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAATGAGGCTGAAAGAAAGCTGGGCTAGAGCGCAAGTTC 2433
Db 2246 TGTCCCAAGCGCTTACAGGCGAGGATCAAGAGAACTAGGGCTAGATCGCGCGCAC 2305
Qy 2434 TACCGGTCAAAATCTCGGTAAACATCACCAAGAAAGACTTCAAGTTCAACGCAAAACGG 2493
Db 2306 CGCCAGTGAAGAACTGTCGCAATATCACTAAAGAGGACCTCAAAATCAACGATGTGACCG 2365
Qy 2494 CAAAATACCGCTCGATCCGAAACCTTCGAGTCTTTGTAGATGCGAAACTCTGCACCT 2553
Db 2366 CACATATTGATGTCAACCTCGAAACCTATAAGGTGAAAGTGGATGCGAAAGAGGTAACT 2425
Qy 2554 CTAACCCACCTCGCAAGTGCCTCTAGCCCGCTACACTTTCTTCTAGG 2604
Db 2426 CTAAAGCAGAGTGAATGAGCTAGCGCACTTTATATTTGTTCTAGG 2476

RESULT 4

US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 755..759
; OTHER INFORMATION: /standard_name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-432-697-19
Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;
Qy 192 ACAAGGAGTAATAGTGAACACTCACCCAAAGAGCAAGAAAGTTCTTTGTTATATTATG 251
Db 29 ATAAGGAGTTTAGGATGAACCTAACGCTTAAGAACTAGACAAGTTAATGCTCCATTATG 88
Qy 252 CGGCGAAGTGGCTAGAAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGCAAGCCA 311
Db 89 CGGCGAGATTGGCAGAGAAGACGCTTGGCGCGTGGTGTGAACCTCAATTACACCGAAGCG 148
Qy 312 TTGCTTACATTTAGTCCCATATTATGACGAGCGCGCTGGGAAAAAAGCTTGGTGGCC 371
Db 149 TCGCGCTCATTTAGCGGCGTGTGATGAAAAGGCGCTGATGTTAATAAGCGTGGCGG 208
Qy 372 AGCTTATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGGGTGA 431
Db 209 ATTGATGCAAGAGGCGAGGACTTGGCTTAAAGAAAGAAATGTGATGACGCGTAGCAA 268
Qy 432 ATATGTTCCCGATCTAGGTGTAGAAGCCACTTTCTGTAGGTGTACGAAACTTTGTAACGT 491
Db 269 GCATGATTCATGAAGTGGGATTTGAAGCTAATCTCCCGATGGAACCAAGCTTTGTAAC 328
Qy 492 TGAATTTGCCCATCGAACCCAGATGACACTTCAAAGCGGCGAAGTGAATTTGGTTGG 551
Db 329 TCCACACTCCGGTAGAGGATATGGCAATTTAGCCCGCGGAGGTCTTCTTAAA--AA 385
Qy 552 ATAAAGACATCGAGTCAATGCAGCAAGAAAGTAAACCGAACTTCAGAGTTACTAATGAAG 611
Db 386 ATGAGGACATTACTATTAAACCGCGCAAGAAAGCCATTAGCTTGAAGTGAAGATAAG 445
Qy 612 GGCTTAAATCCTTGCATGTGGGTAGCCATTTCCACTTTCTTTGAAGCTAACCAAGGCACTAA 671
Db 446 GCGATCGTCTGTGCGAGTGGGATCACATTTCCACTTCTTCGAAGTGAATAAGCTCTTGG 505
Qy 672 AATTCGATCGTGAAGAAAGCCATATGCAACCGCTAGATATTCCTCTCTGCGCAACGCTAC 731
Db 506 ACTTCGATCGCGCAAGAAAGCTTTTGCACAAACGCTAGACATTCGATCTGGAACAGCGGTG 565
Qy 732 GCATTTGGGCGAGGACAAACCCGCAAGGTGAGTTGATTTCTCTGTGTGCGAGTAAAAAAG 791
Db 566 GCTTTGAACCCGGGAGGAAAAAGTGTGGAACCTCATTTGACATCGCGCGGAATAGCGCA 625
Qy 792 TGATTGGCATGAACGGGCTTGTGAATAACATCGCGGATGAACGCCATAACATAAAGCGCG 851
Db 626 TCTATGGCTTTAATTTCTTTTGGTGGATCGCAAGCGGATGCCGATGTTGTAATAAAGCT 685
Qy 852 TTGACAGGCGGAAATCTCACGATTT----- 877
Db 686 TAAACCGCGCTAAAGAAAAAGGTTTGGGTCTGTAACCTGCGGTGTGTGAAGGACATAAG 745
Qy 878 ----ATCAAGTAAGGAGACTCCCATGAAATGAAGAAACAAAGAAATATGTAATACCTACG 933
Db 746 ATAAACAATAAGGAAAAACCATGAAGAAAGATTTTCGAAAGAAATATGTTCTTATGATG 805
Qy 934 GACCCACCAAGGCGGATAAAGTGCCTTTAGGAGATACCGATCTTTGGGCGAGAGTAGAAC 993

Db 806 GTCCCACTACGGGGATCGTGTAGACTCGGCGACACTGATTTGATCTTAGAAGTGAGC 865
Qy 994 ATGACTATACACCTATGGCGAAGAACTTAAATTTGGCGGGGTAAAACTATCGTGAGG 1053
Db 866 ATGATTGCAACACTTATGTTGAAGAGATCAAAATTTGGGGCGGTAAAACTATCGTGATG 925
Qy 1054 GTATGGGTGAGCAATAGCCCTGTATGAACACCCCTAGATTTAGTCACTCACTAACCGGA 1113
Db 926 GGATGAGTCAAAACCAATAGCCCTAGCTCTTATGAATTAGAATTTGGTGTCTCACTAACGCC 985
Qy 1114 TGATTATCGACTACCGGGAATTTACAAGCGACATTTGGATTTAAAGCGCAAAATCC 1173
Db 986 TCATTTGGACTATACGGGCAATTTACAAGCGGACATTTGGGATTTAAAGACGGCAAGATTG 1045
Qy 1174 ATGGCATTTGGCAGGAGGAAACAAGGACATGCAAGATGGCGTAAGCCCTCATATGCTG 1233
Db 1046 CAGGCATTTGCAAGCGAGGCAATAGGACATGCAAGATGGCGTAGATAAATACTTTGGC 1105
Qy 1234 TGGGTGTGGCACAGAACTAGCAGGGAAGGTATGATTTATPACCGCTGGGGGAATCG 1293
Db 1106 TAGGTCCTGTACAGAGGCTTTGGCAGCTGAGGGCTTTGATTTGTAACCGCTGGTGGCATG 1165
Qy 1294 ATTACACACCCACTTCTCTTCCCAACAAATTCCTACCGCTCTAGCCCAATGGCGTTA 1353
Db 1166 ATACGCATATTCACTTTATCTCTCCCAACAAATCCCTACTGCTTTTGGCAGCGGGTTA 1225
Qy 1354 CAACCATTTTGGAGCGGCACAGGCTCTGATGTCGCAAGATGCGCAATGCGCACTACTATCACTC 1413
Db 1226 CAACCATGATTGGAGGAGGCACAGGACCTCGGATGGCAAGATGGCAACCACTCACTC 1285
Qy 1414 CGGCGAAATGGAACTTGCACCGCATTTGCGCGCAGCAGAGAGTATTTCTATGATGTTGG 1473
Db 1286 CCGGACGGCTATCTAAAGATGATGTTGCGTGCAGCGGAAGATACGCCATGATCTAG 1345
Qy 1474 GCTTTTGGCAAGGCAATAGCTCTAGCAAAAAACAACTTTGTAGAACAACTGAAGCGG 1533
Db 1346 GCTTTTGGCTAAGGGGAATGTGTCTTACGAAACCTCTTTACGCGATCAGATTGAAGCAG 1405
Qy 1534 GCGGATTTGTTTAAATTCATGCAAGACTGGGGCAACACCAAGTGGCATCACT 1593
Db 1406 GGGCGATTTGTTTAAATCCACGAAGACTGGGAAGCAACCTTGCAGCTATTCCACACT 1465
Qy 1594 GCTTGAGCGTGGCAGATGAATCAGTGTCAAGTTTGTATCCACACCGATACAGTCAATG 1653
Db 1466 GCCTCAATGTCGCGATGAATAGATGTGCAAGTGGCTATCCACCGATACCTTAAAG 1525
Qy 1654 AGCAGGTTATGTAGATGACACCTTAAATGCAATGCAACGCGCGCCATCCATCCCTACC 1713
Db 1526 AGCGGCTGTGTAGAAGACACCTTAGAGCGGATTTGCGGGCGCACCATCATACCTTCC 1585
Qy 1714 ACATTTGAGGAGCGGTGGAGGACATCACTGATGTTATCACCATGGCAGCGAGCTCA 1773
Db 1586 ACATGAAGGGGTGGGGTGGACACGCTCCAGATGTTATCAAAATGGCAGGGGAATTTA 1645
Qy 1774 ATATTCTACCTCTCCACACCCCACTATTCCTATACCATTAATACGTTTCAGAAC 1833
Db 1646 ACATTTACCGGCTCTACTAACCCGACCAATCTCTTCAACAAAACACTGAAGCCGAGC 1705
Qy 1834 ACTTAGACATGCTCATGACATGCCACCACTAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAATGGTGTGCCACCACTTGTGATAAAGTATCAAGGAGATGTGCAGT 1765
Qy 1894 TTTCTCAAAGCCGTATCCGCCCCGGCTCTATCGCGGTGAAGATGTGCTCATGATAGG 1953
Db 1766 TTGCCGATTCGAGATTTCGCCCCCAAACTATCGCGGTGAAGACCAACTCCATGACATGG 1825
Qy 1954 GTGTGATCGGATGACAGCTCGATTTCGCAAGCAATGGGGCTGCGAGGCAAGTATTC 2013
Db 1826 GGATCTTTTATCACAGCTCCGACTCTCAGGCTATGGACGCGTATGGCGAGGTGATCA 1885
Qy 2014 CTCGAACTTGGCAGACTGCGGATAAGAAATAAAAAAGAAATTTGGTAAGCTTCTCAAGATG 2073
Db 1886 CAGGCATTTGGCAGACGACGACAAACAAAAAGAGTTTGGCGCTTGAAGAGGAA 1945

RESULT 5

US-08-466-248-19

; Sequence 19, Application US/08466248

; Patent No. 6258359

; GENERAL INFORMATION:

; APPLICANT: Labigne, Agnes

; APPLICANT: Sauerbaum, Sebastien

; APPLICANT: Ferrero, Richard L.

; APPLICANT: Thiberge, Jean-Michel

; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST

; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE

; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID

; TITLE OF INVENTION: POLYPEPTIDES

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

; ADDRESSEE: Dunner

; STREET: 1300 I Street, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/466,248

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/447,177

; FILING DATE: 19-MAY-1995

; CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-020000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno
; OTHER INFORMATION: sequence."
; US-08-466-248-19

Query Match 36.9%; Score 1062.6; DB 3; Length 2619;
Best Local Similarity 65.5%; Pred. No. 1e-263;
Matches 1606; Conservative 0; Mismatches 804; Indels 41; Gaps 2;

QY 192 ACAAGGAGTAATAGGTGAACTACACCCCAAGACGAGAAAGCTTAAAGCTCAACCAACCCGAAGCA 311
Db 29 ATAAGGAGTTTAGATGAACTAACGCTTAAAGAACTAGACAAGTTAATGCTTCCATTATG 88

QY 252 CGGCGCAAGTGGCTAGAAAGCGCAAGCAGAGGCGTTAAAGCTCAACCAACCCGAAGCA 311
Db 89 CGGCGAGATTGGCAGAGAACGCTTGGCGCGTGTGTGAACCTCAATTACACGAGCGG 148

QY 312 TTGCTTACATTAGTCCCATATTATGACGAGCGCGCGTGGAAAAAACCCTGTGCC 371
Db 149 TCGCGCTCATTAGCGGCGGTGATGAAAAAGGCGGTGATGTTAATAAAAAGCGTGGCGG 208

QY 372 AGCTTATGAGAGTGTGCATGCACCTTTTGAAGAAAGATGAAGTAATGCCCGGGTGGTA 431
Db 209 ATTTGATGCAAGAAGGCGAGCACTTGGCTTAAAAAAGAAAAATGTGATGGACGCGGTAGCAA 268

QY 432 ATATGGTTCCCGATCTAGGTGTAGAAGCGCACCTTTTCCTGATGGTACGAAACTTGTAACTG 491
Db 269 GCATGATTCATGAGTGGGATTTGAAGCTAATCTCCCGATGGAACCAAGCTTGTAACTA 328

QY 492 TGAATTGGCCCATCGAA CCGATGAGCACATTCAAAGCGGGCGAAAGTGAATTTGGTTGCG 551
Db 329 TCCACACTCCGGTAGAGGATAATGGCAAAATTAGCCCCCGCGAGGTCTTCTTAAA--AA 385

QY 552 ATAAAGATCGAGCTCAATGCAGGCAAGAAAGTAA CCGAACTTGAGGTTACTAATGAAG 611
Db 386 ATGAGGACATTACTATTAAACGCCCGGCAAGAAAGCCATTAGCTTGAAAGTGAATAATAAG 445

QY 612 GGCCTAAATCCTTGCATGTGGTAGCCATTTCCCACTCTTTTGAAGCTAAACAGGCACATAA 671
Db 446 GCGATCGTCTGTGCAGTGGATTCACATTTCCACTCTTTCGAAGTGAATGAAGCTCTTGG 505

QY 672 AATTTCGATCGTAAAAAGCCCTATGGCAAAACGCCCTAGATATTCCCTCTGGCAACACGCTAC 731
Db 506 ACTTTCGATCGCAAAAAGCTTTTGGCAAAACGCCCTAGACATTTGCATCTCGAACACGCGTGC 565

QY 732 GCATTGGGGCAGGACAAACCCGCAAAAGTGCAGTTGATTCCTCTTGGTGGCGAGTAAAAAG 791
Db 566 GCTTTTGAACCCGGGAGGAAAAAGTGTGGAACCTATTGACATCGGCGGGAATAGCGCA 625

```

Qy	792	TGATTGGCATGAACGGCGCTTGTAATAAATCATCGCGATGAACGCCATAAACAATAAAGCGC	851
Db	626	TCTATGGCTTTAATCTTTGGTGGATCGCAAGCCGATCCGATGTGTAAAAAACTCGGCT	685
Qy	852	TTGACAAGCGGAATCTCACGGATT- - - - -	877
Db	686	TAAACCGCTTAAGAAAAAGGTTTTGGGCTGTGTAACCTGCGGTTGTGAAGCGACTAAAG	745
Qy	878	----ATCAAGTAGGAGAGCTCCCATGAAATGAAAAAACAAGAATATGTAATACCTACTACG	933
Db	746	ATAAACAATAAGAAHAACCATGAAAAAGATTTCAGAAAGAATATGTTCTATGTATG	805
Qy	934	GACCCACCAAGCGCGATAAAGTCGCTTAGAGATATACCGATCTTTGGGCGAAGTAGAAC	993
Db	806	GTCCCACTACCGGGGATCGTGTAGACTCGGGACACTGATTTGATCTTTAGAGTGGAGC	865
Qy	994	ATCACTATACCACTATGCGGAGAGACTTTAAATTTTGGCGGGTAAAACTATCCGTTGAGG	1053
Db	866	ATGATTTGCACCACTTATGTGAAGAGATCAAAATTTGGGGCGGTAAAACTATCCGTTGATG	925
Qy	1054	GTATGGGTGAGACCAATAGCCCTGTGTAAGAAAAACAACCCATGATTTTGTGTCATCACTAAACGCCGA	1113
Db	926	GGATGAGTCMAACCAATAGCCCTAGTCTTATGAATTAGATTTGGTGTCTCACTAAACGCC	985
Qy	1114	TGATATTCGACTACACCGGATTTACAAAGCCGACATTCGGGATTTAAAAACGGCAAAATCC	1173
Db	986	TCAATTGTGGACTATACGGGCAATTTACAAAGCCGACATTTGGGATTTAAAGACGGCAAGTTG	1045
Qy	1174	ATGGCATTTGGCAAGGACAGGAACAAGGACATGCAAGATGGCGTTAAGCCCTCATATGGTCG	1233
Db	1046	CAGGCAATGGCAAGGACAGCAATAAGGACATGCAAGATGGCGTAGATAATATCTTTGGC	1105
Qy	1234	TGGGTGTGGGCACAGAAGCACTAGCAGGGGAAAGGTATGATTTATACCGCTGGGGGAATCG	1293
Db	1106	TAGTCTCTGCTACAGAGGCTTTGGCAGCTGAGGGCTTGATTTGAACCGCTGTGGGCATCG	1165
Qy	1294	ATTCACACACCCACTTCTTTCTTCCACAACAATTCCTCACCCTCTAGCCAAATGGCGTTA	1353
Db	1166	ATACGCATATTCACTTTATCTCTCTCCCCACAACATCCCTACTGCTTTTGCAGCGGGTTA	1225
Qy	1354	CAACCAATGTTTGAGCGGGACAGGTCCTGTGTATGGCAAGATGGCATCTATCACTC	1413
Db	1226	CAACCACTGATTTGAGGAGGACAGGACCTGCGGATGGCACGAAATGCGACACCACATCACTC	1285
Qy	1414	CGGGCAATTGGAACTTGCACCGCATGTCGGCGCACGAGAAGTAGTTCATCAATGTGG	1473
Db	1286	CCGGACCGGCTAATCTAAAAAGTATGTCGGTGCAGCGAAGAAATACGCCATGAATCTAG	1345
Qy	1474	GTCTTTTGGCAAGGCAATAGCTCTAGCAAAAAAACAACCTGTAGAACAAAGTAGAAGCGG	1533
Db	1346	GTCTTTTGGCTAAGGGAAATGTCTTTACGAACCTCTTTTACCGGATCAGATTTGAAGCAG	1405
Qy	1534	GCSCGATGTGTTTTAAATTTGCATGAAGAATGGGGCAACAACCAAGTGGGATCGATCACT	1593
Db	1406	GGGCGATTTGGTTTTAAAAATCCACGAAGAATGGGGAAGCACACCTGCAGCTATTCAACCACT	1465
Qy	1594	GCCTTGACGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATG	1653
Db	1466	GCCTCAATGTCGCCGATGAATACGATGTGCAAGTGGCTATCCACACCGATACCCCTTTAACG	1525
Qy	1654	AGGCAGGTTATGTAGATGACACCCTAAATGCAATGAACGGGCGCGCATCCATGCCTACC	1713
Db	1526	AGCGGGCTGTGTAGAAGACACCCCTAGAGCGGATTTCCCGGGCGCACATCCATACCTTCC	1585
Qy	1714	ACATTTGAGGAGCGGGTGGAGGACACTCACTGATGTTTATCACCATGTGGCAGCGGAGCTCA	1773
Db	1586	ACACTGAAGGGCTGGGGGTGGACACGCTCCAGATGTATCAAATGGCAGGGGAATTTA	1645
Qy	1774	ATATTTACCCCTCTCCACCAACCCCACTATTTCCTTATACCAATTAATACGGTTGCAGAAC	1833
Db	1646	ACATTTACCCGCTCTACTAAACCCGACATTCCTTTCCACAAAAACATCTGAAGCCGAGC	1705

```
Qy 1834 ACTTAGACATGCTCATGACATGCAACACCATAGACAAACGATCCGCGAGGATTTACAAT 1893
Db 1706 ACATGACATGTTAAATGGTGTGCCACCATCTGGATAAAAGTATCAAGGAAGATGTGCAGT 1765
Qy 1894 TTTCTCAAAGCGGTATCCGGCCCGGCTCTATCGGGCTGAAGATGTCTCATATATGG 1953
Db 1766 TTGCCGATTTCGAGGATTCGCCCCCAAACTATCGCGCTGAAGACCAAACTCCATGACATGG 1825
Qy 1954 GTGTGATCGGATGACAAGCTCGGATTCGAAGCAATGGGGCGTGCAGGCGAAGTGATTC 2013
Db 1826 GGATCTTTCTATACAGATCTCGACTCTCAGGCTATGGACGGTAGGCGAGTGATCA 1885
Qy 2014 CTCGAACCTTGGCAGACTGCGGATAAGAAATAAAAGAAATTTGGTAAGCTTCTCTGAAGATG 2073
Db 1886 CACGCACTTGGCAGACAGACAGACAAACAAACAAAGAGTTTGGCGCTTGAAGAGGAAA 1945
Qy 2074 GCAAGATTAAGTAATTTCCGATTAAGCGCTACATCTCTCAATATACACTATCAACCCGG 2133
Db 1946 AAGCGGATAACGACAACTTCGCGATCAACGCTACATCTCTAAATACACCATCAACCCCG 2005
Qy 2134 CTTTGACCCAGCGGTGAGCGAGTATATCGGCTCTGTGGAGAGGGCAAGATCGCCGACT 2193
Db 2006 GGATCGCGCATGGATTTCTGACTATGTGGCTCTGTGGAGTGGGCAAAATACGCCGACC 2065
Qy 2194 TGGTGGTGTGGAATCTCTGCTTTTGGCGTAAACCCCAAAATCGTCATCAAAAGCGGTA 2253
Db 2066 TCGTGTCTTGGAGTCCGGCTTTCTTTGGCATTAAAGCCCAATATGATTTAAAGGCGGAT 2125
Qy 2254 TGGTGGTCTTCTGAAATGGGCGATTCTAAACGGTCTGTGCCACTTCCCAACCGGTTT 2313
Db 2126 TTATTTGGCTCTCTCAATGGGCGATGCCAATGCGTCTATTTCCACCCCTCAGCCCGTCT 2185
Qy 2314 ATTACCGGAAATTTTGGGCATCACGGCAAGGGAATTTGACACGAGCATCACTTTTG 2373
Db 2186 ATTACCGTGAATTTTGGACACCATGCGGAAACAAATTCGACACCAATATCACTTTTCG 2245
Qy 2374 TTTCCAAAGTCGCTATGAAATGGCGTGAAAGAAAGCTGGGCTTAGAGCGCCAAAGTTC 2433
Db 2246 TGTCGAGCGGCTTACAGGCGAGGATCAAGAGAACTAGGCTAGATCGCGCGAC 2305
Qy 2434 TACCGGTCAAAATCTGCGGTAACTACATCAACCAAGAAAGACTTTCAAGTTCAACGCAAAACGG 2493
Db 2306 CGCCAGTGAATACTGTGCAATATCACTAAAGAGACCTCAAATTCACGATGTGACCG 2365
Qy 2494 CAAAAATCACCGTGCATCCGAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACT 2553
Db 2366 CACATATTGATGTCAACCCCTGAAACCTTAAAGGTGAAAGTGGATGGCAAGAGGTAACCT 2425
Qy 2554 CTAAACCCACTCGCAAGTGCCTCTAGCCGAGCGCTACACTTTTCTTAGG 2604
Db 2426 CTAAAGCAGCAGATGAATTGAGCTAGCGCAACTTTATAAATTTGTTCTAGG 2476
```

RESULT 6

```
US-09-431-705-1
; Sequence 1, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (2)...(31)
; NAME/KEY: CDS
; LOCATION: (41)...(61)
; NAME/KEY: CDS
; LOCATION: (65)...(799)
; NAME/KEY: CDS
; LOCATION: (803)...(2512)
; NAME/KEY: CDS
; LOCATION: (2516)...(2692)
; NAME/KEY: CDS
; LOCATION: (2696)...(2896)
; NAME/KEY: CDS
; LOCATION: (2900)...(3322)
; NAME/KEY: CDS
; LOCATION: (3326)...(3385)
; NAME/KEY: CDS
; LOCATION: (3389)...(3406)
; NAME/KEY: CDS
; LOCATION: (3410)...(3466)
; NAME/KEY: CDS
; LOCATION: (3470)...(3598)
; NAME/KEY: CDS
; LOCATION: (3602)...(3661)
; NAME/KEY: CDS
; LOCATION: (3665)...(3697)
; NAME/KEY: CDS
; LOCATION: (3701)...(3769)
; NAME/KEY: CDS
; LOCATION: (3773)...(3817)
; NAME/KEY: CDS
; LOCATION: (3821)...(3844)
; NAME/KEY: CDS
; LOCATION: (3848)...(3889)
; US-09-431-705-1
```

```
Query Match 36.7%; Score 1058.6; DB 3; Length 4824;
Best Local Similarity 65.7%; Pred. No. 1.4e-262;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;
```

```
Qy 197 GAGTAATAGGTGAAACTCACACCCAAAGAGCAAGAAAGTTCTTGTATATTATCGGGC 256
Db 77 GAGATCTCCATGAAACTCACCCCAAGAGTTAGATAAGTTGATGCTCCACTACGCTGA 136
Qy 257 GAAGTGGCTAGAAAGCGCAAGAGAGGGCTTAAAGCTCAACCAACCGAAGCCATTGCT 316
Db 137 GAATTTGGCTAAAAACGCAAGAAAGGCAATTAAGCTTAACATATGTAGAGCAGTAGCT 196
Qy 317 TACATTAGTCCCATATTATGACGAGCGCGCGTGAAGAAAAAACCGTTGCCAGCTT 376
Db 197 TTGATTAGTCCCATATTATGGAAGAGAGAGAGCTGGTAAAGAGACTGCGGCTGAATTG 256
Qy 377 ATGGAAGAGTGCATGCACCTTTTGAAGAAAGATGAATGATGCCCGGGTGGGTAAATG 436
Db 257 ATGCAAGAGGGCGGCACTCTTTTAAACCAAGATGATGATGGATGGCGTGGCAAGCATG 316
Qy 437 GTTCCCGATCTAGGTGTAGAAAGCCACCTTTCTCTGATGGTAGCAAACTTTGTAACTGTGAAT 496
Db 317 ATCCATGAAGTGGGTATTGAAGCGATGTTTCTGATGGAGCTAAACTCGTAACCGTGCAT 376
Qy 497 TGGCCCATCGAAACGAGATGAGCATTCAAGGGGGCGAAGTGAATTTGGTTGCCGATAAA 556
Db 377 ACCCCTATTGAGGCAATGGTAAATTAGTTCCTGCTGAGTTG---TTCATAAAAAATGAA 433
Qy 557 GACATCGAGCTCAATGAGGCAAGAGTAACCGAACTTGAGGTTTACTTAATGAAGGCCT 616
Db 434 GACATCACTATCAAGAAAGCAAAAGCCGTTAGCGTGAAGAGTTTAAAAATTTGCGGAC 493
Qy 617 AAATCCTTTGATGTGGGTAGGCATTTCCACTTCTTTGAAGCTTAAACAGGCACTAAATTC 676
Db 494 AGACCGGTTCAATCGGCTCACACTTCCATTTCTTTGAAGTAAGTAGATGCCTAGACTTT 553
```


QY 677 GATCGTGAAGAAAGCCCTATGGCAAGCCCTAGATATATCCCTCTGGCAACACGCTACGCAAT 736
D 678 |||||
D 679 |||||
QY 680 GACAGAGAAAGAACTTTTCGGTAAACGCTTAGACATTTGCGAGCGGACAGCGTAAAGATTT 613
D 681 |||||
QY 682 GGGGAGGACAAACCCGCAAGTGCGAGTTGATTCCTCTTGGTGGGAGTAAAGAAAGTAT 796
D 683 |||||
D 684 GAGCCTGGCGAAGAAAGAAATCCGTTAGAAATTTGATTTGACATTTGGCGGTAAACAGAAATCTTT 673
QY 685 GGCATGAACGGCTTCTGATTAACATCGCGGATGAACGCCATAAACATAAAGCGCTTGAC 856
D 686 |||||
D 687 GATTTAACGCATTTGGTTGATGACAGACAGCAACGAAAGCAAAATTTGCTTTACAC 733
QY 688 AAGGCGAAATCTCACGATTTT-----ATCAAGT 884
D 689 |||||
D 690 AGAGCTAAAGAGCGTGGTTTTTCATGCGCTTAAAGCGATGACAACTATGTAAAGCAAT 793
QY 691 AAGGAGACTCCCATGAAA---ATGAAAAACAAGAAATATGTAATACCTACGAGACCCACC 941
D 692 |||||
D 693 AAGGAGTAAAGAAATGAAAAGAGATTAGCAGAAAGAAATATGTTCTATGATGCTCTACT 853
QY 694 AAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCAAGTGAACATGACTAT 1001
D 695 |||||
D 696 AAGGCGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCAAGTGAACATGACTAT 913
QY 697 ACCACTATGGCGAAGAACTTAAATTTGGCGGGTAAAGTAACTATCCGTGAGGGTATGGGT 1061
D 698 |||||
D 699 ACCATTTATGGCGAAGAGCTTAAATTTGGTGGCGGTAAAGCCCTAAGAGAGCGATGAGC 973
QY 700 CAGAGCAATAGCCCTGATGAAACACCCCTAGATTTAGTCACTAACCGGATGATATC 1121
D 701 |||||
D 702 CAATCTAACAAACCTAGCAAGAGAGTGGATTTAAATTTATCACTAACCGCTTTAAATCGTG 1033
QY 703 GACTACACCGGATTTACAAAGCGGATGGATTTAAAGCGGCAAAATCCATGGCAT 1181
D 704 |||||
D 705 GATTAACCGGATTTTAAAGCGGATTTGGTATTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1093
QY 706 GCAAGGAGGAAACAAAGACATGCAAGATGGCGTAAAGCCCTCATATGGTGGTGGTG 1241
D 707 |||||
D 708 GGTAAAGCGGTAAACAAAGACATGCAAGATGGGTTAAAGCAATCTTAGCGTAGGTCT 1153
QY 709 GGCACAGAGCACTAGCAGGGGAGGTATGATTTATACCGCTGGGGAATCGATTCACAC 1301
D 710 |||||
D 711 GCTACTGAAGCCTTAGCCGGTGAAGTTTGATCGTAAGCGCTGGTGGTATTTGACACACAC 1213
QY 712 ACCACTTCTTCTCCACAACTTCCCTACCGCTTAGCCAAATGGCGTTACACCATG 1361
D 713 |||||
D 714 ATCCACTTCAATTTACCCCAACAAATCCCTACAGCTTTTTCAGAGCGGTGTAACAAACCATG 1273
QY 715 TTGGAGGCGGACAGGCTCTGTAGATGGCAAGTGGCACTACTATCACTCCGGGCAAA 1421
D 716 |||||
D 717 ATTGGTGGGAACCGGCTCTGTGATGGCACTTAATGGCACTACTATCACTCCAGGCGA 1333
QY 718 TGGAACTTGACCGCATGTTGGCGGAGAGAGATTTCTATGAATGTGGGCTTTTG 1481
D 719 |||||
D 720 AGAAATTTAAATGGATGCTCAGAGCGGCTGAAGATATTTCTATGAATTTAGGTTCTTG 1393
QY 721 GGCAGAGCAATAGCTCTAGCAAAACAACTTTGTGAGAACTAGAGAGGCGGCGGAT 1541
D 722 |||||
D 723 GCTAAAGGTAAACGCTTCTAACGATGGAGCTTAGCCGATCAAAATTTGAAGCGGTGGCAT 1453
QY 724 GGTTTTAAATTTGATGAAGACTGGGCAACAAACCAAGTGGATCGATCACTGCTTGAGC 1601
D 725 |||||
D 726 GGTTTTGAATTTCAAGAACTGGGCAACCACTCTCTCTGCAATCAATCAATGCTTAGAT 1513
QY 727 GTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGGATACAGTCAATGAGGAGGT 1661
D 728 |||||
D 729 GTTGGCGAACAATACGATGTGCAAGTGGCTATCGCCACACACACTTTGAATGAAGCGGT 1573
QY 730 TATGTAGATGACACCTTAATGAAGAGCGGCGGCACTCCATGCTTACCAATTTGAG 1721
D 731 |||||
D 732 TGTGTAGAGACACTATGCTGCTATTGCTGGAGCGCACTATGCACACTTTTCCACACTGAA 1633

QY 1722 GGAGCGGTGGAGGACACTCACTGATGTTATACCATGGAGCGAGCTCAATATTTCTA 1781
D 1723 |||||
D 1724 GCGCTGGCGCGGACACGCTCTCTGATATTTAAAGTAGCCGTGAACAACAATTTCTT 1693
QY 1725 CCCTCTCCACCAACCCCACTATTTCCCTATACCAATTAACGTTTACGTTGAGACACTTAGAC 1841
D 1726 |||||
D 1727 CCGCTTCCACTAACCCCACTCCCTTTCCCGTGAATACAGAGCAGAGCATGAGC 1753
QY 1728 ATGCTCATGATGATGCCACCACTAGACAAACGCGATCCGCGAGGATTTTACAAATTTTCTCAA 1901
D 1729 |||||
D 1730 ATGCTTATGTTGGTGGCAGCTTGGATTAAGCAATTAAGAGATGTTCACTGCTGAT 1813
QY 1731 AGCGTATCCGCGCGGCTCTATCGCGCTGAAGATGTCTTCAATGATATGGGTGTGATC 1961
D 1732 |||||
D 1733 TCAAGGATCCGCGCTCAACCAATTTGGCTGAAGACACTTTGCTGATGATGAGTGGGATTTTC 1873
QY 1734 GCGATGACAGCTCGGATTCGCAAGCAATGGGGCGTGCAGCGGAGTATTTCTCGAAT 2021
D 1735 |||||
D 1736 TCAATCAACGATTTGATCTCTCAAGCGATGGCGGTGGGTGAAGTTTATCATTAGAACT 1933
QY 1737 TGGCAGACTCGGATAAGATTAAGAAAGATTTTGGTAAAGCTTCTCTGAAGATGGCAAGAT 2081
D 1738 |||||
D 1739 TGGCAACAGCTGACAAAACAGAAAGATTTGGCGCTTGAAGAGAGAAAGAGCGAT 1993
QY 1740 AACGATATTTCCGCTATTAAGCGCTACATCTTCAAAATACACTATCAACCCCGCTTTGACC 2141
D 1741 |||||
D 1742 AACGCAACTTTCAGGATCAAGCGCTACTTGTCTAAATACACATTAACCCAGCGATCGCT 2053
QY 1743 CAGCGGTGAGCGATATATCGGCTCTGTGGAAGAGGCAAGATCGCGCACTTGTGTGGTG 2201
D 1744 |||||
D 1745 CATGGGATTAAGCGATGATGAGTTTCAAGTGAAGTGGGCAAGTGGCTGACTTGGTATTTG 2113
QY 1746 TGGAACTCTGCTTTTGGCGTAAACCCAAATCGTGATCAAGGCGGTATGTGGTG 2261
D 1747 |||||
D 1748 TGGAGTCCAGATTTCTTTGGCGTAAACCCAAATGATCATCAAGGCGGATTCATTGG 2173
QY 1749 TTCTCTGAAATGGCGGATTTCTAACCGCTCTGTGCGCACTCCCGCAACCGGTTTATACCGC 2321
D 1750 |||||
D 1751 TTAAGCCAAATGGCGATGCGAACGCTTCTATCCCTACCCCAACCGGTTTATTACAGA 2233
QY 1752 GAAATGTTGGGCAATCGGCAAGGCAAAATTTGACACGAGCATCACTTTTGTGTTCCAAA 2381
D 1753 |||||
D 1754 GAAATGTTGCTCATCATGTGTAAGCTAAATACGATGCAAAACATCACTTTTGTGCTCAA 2293
QY 1755 GTGCGCTATGAAATGGCTGAAAGAAAGCTGGGCTTAGAGCGCAAGTCTACCGGTC 2441
D 1756 |||||
D 1757 GCGGCTTATGCAAGGCAATTAAGAGAGATTAAGGACTTTGAAGACAAAGTGTTCGCGGTA 2353
QY 1758 AAAAACTGGCGTAAACATCAACCAAGAAAGCTTCAAGTTCAACGACAAACCGCAAAATC 2501
D 1759 |||||
D 1760 AAAAACTGGCAAAATATCACTTAAAGAGACATGCAATTCACGACACTACCGCTCACATT 2413
QY 1761 ACCGTGCAATCGAAAACTTGGAGGTCTTTGTAGATGCAAACTCTGCACCTCTAAACCC 2561
D 1762 |||||
D 1763 GAAGTCAATCTGAACTTACCATGTGTTGCTGATGGCAAGAGTAACCTTCTAAACCA 2473
QY 1764 ACCTGCAAGTGCCTCTAGCCAGCGCTACACTTTCTTCTAGG 2604
D 1765 |||||
D 1766 GCAATAAAGTGAAGTGGCGCAACTCTTTAGCATTTTCTAGG 2516

RESULT 7

US-09-431-705-19
; Sequence 19, Application US/09431705
; Patent No. 6585975
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431,705


```
; CURRENT FILING DATE: 1999-11-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
; Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (3893)...(3934)
; NAME/KEY: CDS
; LOCATION: (3938)...(4027)
; NAME/KEY: CDS
; LOCATION: (4031)...(4285)
; NAME/KEY: CDS
; LOCATION: (4289)...(4300)
; NAME/KEY: CDS
; LOCATION: (4304)...(4408)
; NAME/KEY: CDS
; LOCATION: (4412)...(4471)
; NAME/KEY: CDS
; LOCATION: (4475)...(4588)
; NAME/KEY: CDS
; LOCATION: (4592)...(4669)
; NAME/KEY: CDS
; LOCATION: (4673)...(4711)
; NAME/KEY: CDS
; LOCATION: (4715)...(4774)
; NAME/KEY: CDS
; LOCATION: (4784)...(4824)
; US-09-431-705-19

Query Match      36.7%; Score 1058.6; DB 3; Length 4824;
Best Local Similarity 65.7%; Pred. No. 1.4e-262;
Matches 1606; Conservative 0; Mismatches 799; Indels 38; Gaps 3;

QY 197 GAGTAATAGTGAACCTCACACCAAGAGCAGAGAGTTCTTGTATATATATGCGGC 256
DB 77 GAGATCTCCATGAACCTCACCCCAAGAGTTAGATGAAGTTGATGCTCCACTACGCTGGA 136
QY 257 GAAGTGGCTAGAAAGCGCAAGCAGAGGCTTAAAGCTCAACCAACCCGAAGCATTGCT 316
DB 137 GAATGGCTTAAAGCAAGAAAGAGCAATTAAGCTTAATGATGAGAGCAGTAGCT 196
QY 317 TACATTAGTCCCATATTATGACGAGCGCGCTGGAAAAAACCCTTGGCCAGCTT 376
DB 197 TTGATTAGTCCCATATTATGGAAGAGCGAGAGCTGGTAAAAAGACTGCGGCTGAATTG 256
QY 377 ATGGAAGAGTGCACTGCTTTTGAAGAAAGATGAAGTAATGCCGGGTGGTAAATG 436
DB 257 ATGCAAGAGGGCGCATCTCTTTTAAACCCAGATGATGTGATGGCGGTGGCAAGCATG 316
QY 437 GTTCCCGATCTAGTGTAGAACCCACTTCTCTGATGGTACGAACTTGTAACTGTGAAT 496
DB 317 ATCCATGAAGTGGGTATTGAAGCGATGTTTCTGATGGGACTAACTCGTAAACCGTGAT 376
QY 497 TGGCCCATCGAACAGATGAGCACTTCAAGCGGGCGAAGTGAATTTGGTTGCGATAAA 556
DB 377 ACCCCTATTGAGGCGCAATGTAATTAGTTCTCTGGTGAGTTG--TTCCTTAAAAATGAA 433
QY 557 GACATCGAGCTCAATGAGCGCAAGATTAACCGAACTTGAGGTACTAATGAAGGCGCT 616
DB 434 GACATCACTATCAACGAAGCAAAAAAGCGGTTAGCGTGAAGTTAAAAATGTTGGCGAC 493
QY 617 AATCTCTGCATGGGTAGCTATTTCCACTCTTTGAAGCTTAACAGGCACTAAATTC 676
DB 494 AGACCGGTTCAATTCGGCTCACACTTCCATTTCTTTGAAGTGAATAGATGCCCTAGACTTT 553
QY 677 GATCTGTAAGAACCTATGCAACCGCTAGATATTCCCTCTGGCAACAGCTACGCATT 736
DB 554 GACAGAGAAAAAATCTTCGGTAAACGCTTAGACATTGCGAGCGGACAGCGGTAAAGATT 613

737 GGGCAGGACAAACCCGCAAAAGTGAGTTGATCTCTTGTGGCAGTAAAAAGTGATT 796
614 GAGCTGGCGAAGAAAAATCCGTAGAAATTGATTGACATTTGGCGGTAAACAGAGAATCTTT 673
797 GGCATGAACGGCTTGTGAATAACATCGCGGATGAACGCCATAACATAAAGCGCTTGAC 856
674 GGATTTAACGCATTGGTTGTAGACAAGACAGACAACGAAGCAAAAAAATTTGCTTTACAC 733
857 AAGCGCAAAATCTCACGGATT-----ATCAAGT 884
734 AGAGCTAAGAGCGTGGTTTCATGGCGCTAAAGCGATGACAACTATGTAAAAACAATT 793
885 AAGGAGACTCCCATGAAA---ATGAAAAAACAAGAAATATGTAATACTACGACCCACC 941
794 AAGGAGTAAGAAAATGAAAAAGATTAGCAGAAAAAGAAATATGTTCTATGTATGCTCTACT 853
942 AAGCGGATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAGATAGAAATGACTAT 1001
854 ACAGCGATAAAGTGAGATTGGCGGATACAGATTTGATCGCTGAAATAGAAATGACTATAC 913
1002 ACCACTATGCGGAAGAACTTAAATTTGCGCGGCTAAACCTATCCGTGAGGTATGGGT 1061
914 ACCATTTATGCGGAAGAGCTTAAATTCGTGCGGTAAACCTTAAGAGAAGGCATGAGC 973
1062 CAGAGCAATAGCCCTGATGAAAAACACCCCTAGATTAGTATCATCACTAAACGCGATGATTATC 1121
974 CAATCTAACACCTTAGCAAGAGAGTTGGATTATTAATTCACATAACGCTTAAATCGTG 1033
1122 GACTACACCGGATTTACAAAGCGGACATTTGGGATTAATAACGCGCAAAATTCATGCCATT 1181
1034 GATTACACCGGATTTATAAAGCGGATATGGTATTAAGATGGCAAAATTCGCTGGCAAT 1093
1182 GGCAGGACGAAACAGGACATCAAGATGGGTAAAGCCCTCATATGTCGTGGGTGGTG 1241
1094 GGTAAAGCGGTAAACAAAGACATGCAAGATGGCGTTAAACAAATCTTAGCGTAGTCTCT 1153
1242 GGCAAGAGCAGCTAGCAGGGGAGGTATGATTATTACCGCTGGGGGAATCCGATTACACAC 1301
1154 GCTACTAGAGCTTAGCGGTGAAGTTGATCGTAAACGCTGGTGGTATTCACACACAC 1213
1302 ACCACTTCTCTTCCACAAATTCCTACCGCTTAGCOATGGCGTTACACCATG 1361
1214 ATCCACTTCTTTCACCCCAAAATCCTACAGCTTTTGAAGCGGTGTAACCAACATG 1273
1362 TTGGAGCGGCACAGGCTCTGTAGATGGCAGATGCGACTACTATCACTCCGGGCAAA 1421
1274 ATTGGTGGTGAACCGGCTCTGCTGATGGCACTAATGCGACTACTATCACTCCAGGCAGA 1333
1422 TGGAATTTGACCCGCTGTCGCGCAGCAGAGAGTATTCATGAATGTGGGCTTTTGTG 1481
1334 AGAAATTTAAATGGATGCTCAGAGCGGCTGAAGAAATATTCATGAAATTTAGTTTCTTG 1393
1482 GGCAAAAGCAATAGCTCTTAGCAAAAAAACAATTTGTAGAACAAAGTAGAAGCGGCGGATT 1541
1394 GCTAAAGGTAAAGCTTTAAACGATCGAGCTTAGCCGATCAAAATGAAGCGGCTGCGATT 1453
1542 GGTTTTAAATTTGATGAAGACTTGGGGCAACAACAAGTGGGATCGATCACTGCTTGAGC 1601
1454 GGCTTTGCAATTCAGAAAGACTTGGGGCAACCACTCTCTTCTGCAATCAATCATCGTTAGAT 1513
1602 GTGCGCATGAATAGATGTGCAAGTTTGTATCCACCGATACAGTCAATGAGCAGCT 1661
1514 GTTCCGACAAATACGATGTGCAAGTCCGCTATCGCCACAGACACTTTGAATGAAGCGGT 1573
1662 TATGTAGATGACACCTTAAATGAAGCAATGAAGCGGCGGCGCATCCATGCTCTACCAATTGAG 1721
1574 TGTGTAGAGACACTATGGCTGCTATTGCTGAGCGCACTATGCACTTTCCACACTGAA 1633
1722 GGAGCGGTGGAGGACACTCACTCTGATGTTATCAACCATGGCAGGCGAGCTCAATATCTTA 1781
1634 GGCCTGGCGGCGACACGCTCCTCTGATATTATTAAGTAGCGGTGAACACAACTCTT 1693
```

QY 1782 CCTCTCTCACACCCCGGCTATTCCTATACCATTAATACGGTTGCAGAACACTTAGAC 1841
Db 1694 CCGCTTCCACTAACCCCACTCCCTTTCCCGTGAATACAGAACGAGCAGCATGGAC 1753
QY 1842 ATGCTCATGACATGCCACACCTAGACAAACGATCCGCGAGGATTTACAAATTTCTCAA 1901
Db 1754 ATGCTTATGCTGCGACCACTTTGGATAAAGCATTAAGAGAGATGTTGAGTTGCGTGAT 1813
QY 1902 AGCCGATCCGCGCGGCTCTATCGCGGCTGAAGATGTCTCCATGATATGGTGATC 1961
Db 1814 TCAAGGATCCGCGCTCAACACCTTGGCTGAAGACACTTTGATGACATGGGATTTTC 1873
QY 1962 GCGATGACAGCTCGATTCGACAAATGGGGCTGCGAGGCAAGTGAATCTCGAACT 2021
Db 1874 TCAATCACCAGTTCTGACTCTCAAGCGATGGGCGGTGGGTGAAGTTATCACTAGAACT 1933
QY 2022 TGGGAGCTGCGGATAAGATAAAGAAATTTGGTAAGCTTCTCAAGATGSCAAGAT 2081
Db 1934 TGGCAACAGCTGACAAAACCAAGAAAGATTTGGCGGCTTGAAGAGAAAGAGCGGAT 1993
QY 2082 AACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATACACTATCAACCCCGCTTTGACC 2141
Db 1994 AACGACAATTCAGGATCAACCGCTACTTGTCTAATACACCATTAACCCAGCGATCGCT 2053
QY 2142 CAGCGGCTGAGGAGTATATCGGCTCTGTGGAGAGGGCAAGATCCCGACTTTGGTGGTG 2201
Db 2054 CATGGGATTAGCGAGTATGTAGGTTTCAGTAGAAGTGGGCAAGTGGCTGACTTGGTATTG 2113
QY 2202 TGGATCCTGCGCTTTTTCGCGGTAAACCCCAAAATCGTATCAAGCGGCTATGGTGGTC 2261
Db 2114 TGGAGTCCAGCATCTTTGGCGTGAACCCCAACATGATCATCAAGCGGATTCATTTGG 2173
QY 2262 TTCTCTGAAATGGCGGATTTAAACCGGCTGTGCGCCACTCCCCAACCGGTTTATTAACCGC 2321
Db 2174 TTAAGCAATGGCGATGGAAGCTTCTATCCCTACCCCAACCGGTTTATTACAGA 2233
QY 2322 GAAATGTTGGGATCACCGCAAGGCAATTTGACACGAGCATCTTTTGTTCAAA 2381
Db 2234 GAAATGTTGCTCATCATGTAAGCTAAATACGATGCAAAATCATCTTTTGTGTCAA 2293
QY 2382 GTGCGCTATGAAATGGCTGAAGAAAGCTGGCTTAGAGCGGCTTACCGGCTAC 2441
Db 2294 GCGGCTTATGACAAAGGCAATTAAGAGAAATTAGGACTTTGAAAGACAAGTTTGGCGGTA 2353
QY 2442 AAAAAGCTCGGTAACATACCAAGAAAGACTTCAAGTTCAACGACAAAAACGCAAAATC 2501
Db 2354 AAAAATTGAGAAATATCACTAAAAAGACATGCAATTAACGACACTACCGCTCACAT 2413
QY 2502 ACCGTCGATCCGAAAACTTCGAGGTCTTTGTAGATGGCAAACTCTGCACTCTAAACCC 2561
Db 2414 GAAGTCAATCTGAAACTTACCATGTGTTGCTGATGGCAAGAAAGTAATCTTAAACCA 2473
QY 2562 ACCTGCAAGTGCTCTAGCCGAGCGGTACACTTTCTTTAGG 2604
Db 2474 GCCATAAAGTGAAGTTGGCGCAACTCTTTAGCATTTTCTAGG 2516

RESULT 8
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frishauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA

ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72mb
COMPUTER: IBM PC compatible (NEC PC-9801 RX)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07732,242C
FILING DATE: 19910718
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPN 2-210178
FILING DATE: 10-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910532/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-732-242C-8
Query Match 25.6%; Score 737.6; DB 2; Length 6131;
Best Local Similarity 58.3%; Pred. No. 1.1e-179;
Matches 1442; Conservative 0; Mismatches 969; Indels 63; Gaps 6;
QY 185 ATTTACACAGAGAGTAAATAGGTGAACCTCACACCCAAAGAGCAAGAAAGTTCTTGTTA 244
Db 317 ACTTAAATAGGAGGTATCCATGAAACTGACTGAAATGGAAGAGCTCATGATT 376
QY 245 TATTATGCGGCGAAGTGGCTAGAAAGCGCAAGAGAGGCTTAAAGCTCAACCAACCC 304
Db 377 GTAGTGGCGGCTGACTTGGCCCGCTGTAAGAGCGGGCTTAAATTAATATATCT 436
QY 305 GAAGCCATTTCTTACATAGTGTCCCATATATTAGCAGAAAGCGCGCTGGAAAAAACC 364
Db 437 GAAGCTGTGCAATGATTACATATGAAGTGTGGAGGGCGCGGGATGG---AAAAAG 493
QY 365 GTTGGCCAGCTTATGGAAGAGTGCATGCATTTTGTGAAAAAAGATGAAGTAAATGCCCGG 424
Db 494 GTAGCTCAGTTAATGCAATACGGTGCAACGATTTCTTACAAAAAGAGATGTAATGGAAGG 553
QY 425 GTGGTAATATGGTTCCCGGATCTAGGTGTAGAGCCACCTTTCTCTGATGGTACGAACTT 484
Db 554 GTTGGCAAAATGATCCCGGATATTTCAAATTTAGGCAACCTTTCTCTGATGGAAACAAAGCTT 613
QY 485 GTAACTGTGAATTTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGGCGAAAGTGAATTT 544
Db 614 GTCAOGTTTCATGACCCGATCCGTTAATGAGGAGGACGTACGATGATACAGGGAGTA 673
QY 545 GGTTCGCAATAAGACATCGAGCTCAATGCGAGGCAAGAAAGTAAACCGAATCTGAGGTACT 604
Db 674 TGTATTAAAAAAGAACCTATTTTATGCAATCAAAATAAGCAGACGATCAAGATTCGCGGT 733
QY 605 AATGAA----GGGCTAAATCCTTGCATGTGGGTAGCCATTTCCACTTCTTTGAGCTAA 660
Db 734 GTTAAACCGGGCGGATCGACTGTTGAGGTGGTTCCTCATTTTTCATTTTGAAGTAA 793
QY 661 CAAGGCACATAAATTCGATCGTGAAGAAAGCTATGGCAAGCGCTAGATATATTCCTCTGG 720
Db 794 TCAATCGCTTCAATTTTCATCTGTGAAAAAGCAATTTGGCATGCGTTTGAATATTCGCGTGG 853
QY 721 CAACACGCTACGCAATTTGGGCGAGGACAAACCCGCAAGTGCAGTTGATTCCTCTTTGGTGG 780
Db 854 AACGGCGTTTCGCTTCGAGCCCGGAGATGCGAAAGAGTAGAAATAATTTCCATTTTCAGG 913

Qy	781	CAGTAAAAAAGTGTGATTCGGCATGAACGGGCTTTGTGAATAACATCGCGATGAACGCCATAA	840
Db	914	TGAACGCAAAAGTGTATGGTTTAAATAATGTAAACGA-----	948
Qy	841	ACATAAAGCGCTTGACAAGCGGAATCTCACGGATTTATCAAGTAAGGAGACTCCCAATGA	900
Db	949	--ATGGATCAGTTTCAAAATGGGAAAAGAAAATGAGTT-----	985
Qy	901	AAATGAAAAACAAGATATGTAAATACCTACGACCCACCAAAAGGCGATAAAGTGCCT	960
Db	987	CGATGTCTCGAAAAGCAATATCGCGATATGTTTGGACCAACTGTGCGGCGACGCATTCGTT	1046
Qy	961	TAGGAGATACCGATCTTTTGGSCAGAAGTAGAACATGACTATACCACTATGGCGCAAGAAC	1020
Db	1047	TGGCAGATTCAGAAATGTTTATCGAAATGAAAAGGACTATACAACTATGGAGATGAGG	1106
Qy	1021	TTAAATTTGGCGGGGTAAAACTATCCGTGAGGGTATGGGTTCAGAGCAAT---AGCCCTG	1077
Db	1107	TAAAGTTTGGCGCGCAAGGTGATCCGAGATGGAATGGGAGCATCTTTGGCGCAAA	1166
Qy	1078	ATGAAAACACCCTAGATTATGTCATCACTAAACCGGATGATTTATCGACTACACCGGAGTTT	1137
Db	1167	CGCATGAATCGTTCGATCTCGTATTAACAAATGCGATTATTTGTGTGATTAACACAGGTATTT	1226
Qy	1138	ACAAGCCGACATTTGGGATTTAAAAACCGCAAAATCCATGCGCATTTGGCAAGCGCAGGAACA	1197
Db	1227	ATAAAGCAGATATCGGCATATAAAGATGGAATGATTTGCCTCCATAGGAAAAGCGGGGAACC	1286
Qy	1198	AGGACATGCAAGATGGCGTGAAGCCCTCATATGGTGTGGGTGTGGGCACAGAAGCACTAG	1257
Db	1287	CGTTGTTAATGACGGGGTCG-----ATATGGTGATTTGAGCAGCAAGAGTCAATAG	1340
Qy	1258	CAGGGGAAGGTATGATTAATTACCGCTGGGGGAATTCGATTACACACCCACTTCCTTTCTC	1317
Db	1341	CCGCAGAAGGGATGATTTGTACACGCGGAGGAATAGATGCTCATATTTCACTTTATTTGCC	1400
Qy	1318	CACAACAATTCCTTACCGCTCTAGCCAAATGGCGTTTAAACCATGTTTGGAGGCGGCACAG	1377
Db	1401	CTCAGCAAAATCGAAAACCGCTCTTGCAATCGGGTGTGACCACTATGATTTGGCGGAGGAACAG	1460
Qy	1378	GTCTGTAGATGGCACGAATTCGCACTACTATCACTCCGGCCAAATGGAATTCGCACCGCA	1437
Db	1461	GACCCGCTCAGGCACAAATGCCACTACTTTGTACACCGGGGCCCTTGGAAATATCCATCGTA	1520
Qy	1438	TGTTGCGCGCAGCAGAAGAGTATTTCTATGAATGTGGCTTTTTTGGGCAAAAGGCAATAGCT	1497
Db	1521	TGCTTCAAGCAGCCGAAGAAATCCCGATAAACTTTGGGCTTTTTAGGAAAGGGAACCTGTT	1580
Qy	1498	CTAGCAAAAAACAACCTTGTAGAAACAAGTAGAAGCGGCGCGGATTTGGTTTTAAATTTGCATG	1557
Db	1581	CAGATGAGGCTCCTTTTAAAGGAACAAATTGAAGCGGAGCGGTGGGATTAAGCTTCAACG	1640
Qy	1558	AAGACTGGGCACAACCAAGTCCGATTCGATCACTGCTTGAGCGTGGCAGATGAATACG	1617
Db	1641	AAGATTTGGGATCGACGCGCGCGCTATTGATACATCTTTGAAAGTGGCGGATCGATATG	1700
Qy	1618	ATGTCAAGTTTTGTATCCACACCGATACAGTCAATGAGGCAGGTATATGTAGATGACACCC	1677
Db	1701	ATGTCAAGTAGCGATTTATACAGACACTTTTAAATGAAGCGGATTTGTTCGAGGATACCTT	1760
Qy	1678	TAAATGCAATGAACGGGCGGCCATCCATGCCTACCACTTGAGGGAGCGGTGGAGGAC	1737
Db	1761	TGAAGCCATAGACGGTTCGAGTGATTTATACCTATCATACAAAGGGGCTTGGCGGGGAC	1820
Qy	1738	ACTCACCTGATTTATCCCATGCGAGCGAGCTCAATATTTCTACCTCTCCACCAACCC	1797
Db	1821	ATGCTCCGATATTTATAAAGCGGCGGCTTCCCGAATATTTTGCCTTCTTCCACGAATC	1880
Qy	1798	CCACTATTCCTCATACCAATTAATACGGTTGAGAAACACTTAGACATGCTCATGACATGCC	1857
Db	1881	CAACTCCACCTTATACTATCAATCACTTTTGAAGAGCATTTAGATATGTTAATGTTTGGC	1940

RESULT 9
US-08-967-513-1
; Sequence 1, Application US/08967513
; Patent No. 5783436
; GENERAL INFORMATION:
; APPLICANT: Robert P. Hausinger
; TITLE OF INVENTION: Mutant Urease and Method
; TITLE OF INVENTION: of Use For Determination
; TITLE OF INVENTION: of Urea
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ian C. McLeod
; STREET: 2190 Commons Parkway
; CITY: Okemos
; STATE: Michigan
; COUNTRY: USA

```

/ ZIP: 48864
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
/ MEDIUM TYPE: storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: Wordperfect 5.1
/ CURRENT APPLICATION DATA:
/ FILING DATE: 11-NOV-1997
/ FILING DATE: 11-NOV-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/687,645
/ FILING DATE: July 26, 1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ian C. McLeod
/ REGISTRATION NUMBER: 20,931
/ REFERENCE/DOCKET NUMBER: MSU 4.1-309
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (517) 347-4100
/ TELEFAX: (517) 347-4103
/ TELEX: No. 5783436e
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2400
/ TYPE: nucleotides
/ STRANDEDNESS: Single
/ TOPOLOGY: Linear
/ MOLECULE TYPE:
/ DESCRIPTION: cDNA
/ HYPOTHEICAL: No
/ ANTI-SENSE: No
/ ORIGINAL SOURCE:
/ ORGANISM: Klebsiella aerogenes
/ STRAIN: CG253
/ INDIVIDUAL ISOLATE:
/ CELL TYPE: N/A
/ FEATURE:
/ NAME/KEY: cDNA encoding mutant urease
/ NAME/KEY: 'H219Q
/ LOCATION: Mutificication at position 1312 to
/ LOCATION: glutamine
/ IDENTIFICATION METHOD: Sequencing
/ OTHER INFORMATION:
/ US-08-967-513-1

Query Match 24.1%; Score 693.4; DB 2; Length 2400;
Best Local Similarity 57.1%; Pred. No. 1.7e-168;
Matches 1385; Conservative 0; Mismatches 971; Indels 69; Gaps 4;

QY 195 AGGAGTAAATAGGTGAAACTCACACCCAAAGACGAAAGAAAGTTCTTGTATATTATGCGG 254
DB 20 AAGAGAAGTATGGAAGTACGCCCGGAGAAAGACAAAGCTGTGCTGTTTACGCGG 79
QY 255 GCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAAGCCATTG 314
DB 80 CGCTGGTGGCGGAGCGTCGCTGCGCGCGCTGAAGCTCAACTATCCGGAGTCCGCTGG 139
QY 315 CTATACATTAGTCCCATATTATGAGCAAGCGCGCGCTGGAAGAAAGAAACCCGTTGCCAGC 374
DB 140 CCCTGATCAGCGCCCTTTATTATGGAAGGCGCTCGGGACCG---CAAAAGCGTGGCCTCGC 196
QY 375 TTATGGAAGAGTGCATCAGCTTTTGTGAAAGATGAAGTAATGCCCGGGTGGGTAATA 434
DB 197 TGATGGAGGAGGCGGTGACAGTCTGACCCCGGAGAGGATGAGAGGGCGTCCCGGAAA 256
QY 435 TGGTTCCCGCATAGTGGTGTAGAGCCACCTTTTCTGATGTGTAGCAAACTTGTAACTGTGA 494
DB 257 TGATCCCGGATATCCAGGTGAGAGCCACCTTCCCGGACGCTCGAAGCTGTCACCGTTC 316
QY 495 ATTGGGCCATCGAACACGATGAGCATTTCAAAGCGGCGGAAGTGAATTTGGTTGGGATA 554
DB 317 ACAACCCGATATCTGAGGTAGGCGCATGATCCCGGTGAATATCAGTTTAAGCCCGGTC 376

```


QY	375	TTATGGAAGAGTGCATGCACTTTTGGAAAAAGATGAAGTAATGCCCCGGGTGGGTAATA	434
Db	197	TCATGGAGGAAGCCGCTCAGCTCTGACCCGGAGCAGGTGATGGAGGCGCTCCCGGAAA	256
QY	435	TGGTTCCTCCGATCAGTGTGAGAAGCACCTTCTCTGATGAGTGAAGAACTGTAACTGTGA	494
Db	257	TGATCCCGGATATCCAGGTTCGAAGCCACCTTCCCGAGCGCTCGAAGCTGGTCACCGTTC	316
QY	495	ATTGGCCCATCGAAACAGATGAGCACTTCAAAGCGGCGGAAGTGAATTTGGTTGGGATA	554
Db	317	ACAAACCGGATTTATCAGGTAGCGGCATGATCCCCGGTGAATATCAGTTAAGCCCGTTC	376
QY	555	AAGACATCAGCTCAATGCAAGCAAGAAAGTAAACCGAACTTGAGGTTCCTAATGAAGGGC	614
Db	377	---AGATAGCCCTGAATACCGGCGGCGCAACCTGTGCGGTGGTTCGTTTGAGAACCAACGGCG	433
QY	615	CTAAATCTTGTGATGTTGGGTAGCCATTTCCACTTTCTTTGAGCTAAACAGGCACTAAAT	674
Db	434	ATCGGCCGATTCAGGTTCGGTTCGCACTACCAATTTCCCGGAGGTAAACCGCGCTGAAGT	493
QY	675	TCGATCGTGAAGAGCTTATGGCAAGCGCTAGATATTCCTCTGTCGCAACACGCTACGCA	734
Db	494	TCGACCGTCAAGCGCGCGCTATCGCTGTAATATCCCGGCGGCAACGCGGTACGCT	553
QY	735	TTGGGCGAGGCAAAACCCGAAAGTGCAGTTGATTCCTCTTGTGTGCAGTAAAGAAAGTGA	794
Db	554	TTGAACCCGCGCAGAAACGGAGTGCAGTGTGTGCTTCGCGGCTCACCGCGCTCT	613
QY	795	TTGGCATGAACGGGCTTTGTAATPAACATCGCGATGAACGCCATAAACATAAAGCGCTTG	854
Db	614	TCGGCTTCGCGCGAGGTCAATGGGCGCTCTGGAGGTAAACGAT-----	657
QY	855	ACAAGCGAAATCTCAGCGATTTATCAAGTAAGGAGCTCCCATGAAATGAAGAAACAA	914
Db	658	-----GAGTAATATTTCAACGCCAG	676
QY	915	GAATATGTAATACCTTACGGACCCCAAGGCGATAAAGTGGCTTAGAGATACCGAT	974
Db	677	GCCTATGCCGATATGTTCCGGCCCAACGCTCGGCGACAGGTGCGCTTGCAGATACGAG	736
QY	975	CTTTGGCGAGAGTAGAAATGACTATACACTATGGCGAAGAACTTAAATTTGGCGCG	1034
Db	737	CTGTGATCGAGTGGAGGACGATTTGACCACTACGGGAGAGGTCAAATTTCCGCGCG	796
QY	1035	GGTAACACTATCGGTGAGGTTAGGTCAGAGCAATAGCCCTCATGAAGAACCCCTAGAT	1094
Db	797	GGCAAAAGTATCCGACCGCATGGGCGAGGACAGATGCTGGCCCGCACTGTGTGAC	856
QY	1095	TTAGTCATCACTAAACCGGATGATTATCGACTACACCGGGATTTACAAAGCGGACATTTGG	1154
Db	857	CTGGTGCTCACCAACCGTGTGATCGTGCATCTGCGGGATCGTTAAGGCCGATATCGGC	916
QY	1155	ATTAAACCGGAAAAATCCATGSCATTTGGCAAGCAGGAAACAAAGACATGCAAGATGGC	1214
Db	917	GTGAAGGACCGCCGGATCTTCGCCATCGGCAAGCGCGCAACCCCGACATCCAGCCCAAC	976
QY	1215	GTAAGCCCTCATATGTTGTTGGTGTGGGACAGAGACATAGCAGGGAAGGTATGATT	1274
Db	977	GTCAAC-----ATCCCCATCGCGGTGCGACGGAAGTATGATCGCCCGCGAAGAAAAAT	1030
QY	1275	ATTACCGCTGGGGAAATCGATTCACACCCACTTCTTCTTCTCCACAGCAATTTCCCTACC	1334
Db	1031	GTACCCCGCGCGGATCGATACCCATATTCATGGAATCTGTCGCGAGCGCGGGAAGAG	1090
QY	1335	GCTCTAGCAATGGCGTTACAAACCTATTTGAGAGGGGCGACAGGTCTCTGATAGGCACG	1394
Db	1091	GGCTGTCTCTGGCGTGACCAATATGTTGGCGGGCGACCGCGCGCGCGCGGACCC	1150
QY	1395	AATGCGACTACTATCATCTCGGGGAAATGAAATTGACCGCATGTTGCGCGCAGCAGAA	1454
Db	1151	CATGCCAACCACTGCACCCCGGCGCGTGTATATCTACGGCATGTGCGAGGCGGCGAC	1210
QY	1455	GAGTATTTCTATGAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATT	1514
Db	1211	AGCTCCCGCTCAATATCGGCCTCTGGGCAAGGAAACGTTTCTCAGCCGATGCCCCTG	1270
QY	1515	GTAGAACAGTAGAAAGCGGCGGATTTTAAATTTGATGAAGACTGGGGGACACACA	1574
Db	1271	CGGAGCAGGTGGCGGAGCGCTTATTTGGCTGAAGATCCAAGAGGACTGGGGCGCCACC	1330
QY	1575	CCAAGTGCATTCGATCACTGCTTGAAGCTGGCAGATGAATACGATGTGCAAGTTTGTATC	1634
Db	1331	CGGGCGGATCGACTGTGGTTAAACGTCGCCGATGAATGGACATCCAGGTGCGCCCTG	1390
QY	1635	CACACGATACAGTCAATAGGCGAGGTATGTAGATGACACCTTAATGCAATGAACGGG	1694
Db	1391	CACAGCGACACCTGATGAATCCGCTTTGTGGAAGACACCTCGCGCCCATCGGCGGG	1450
QY	1695	CGGCGCATCCATCGCTTACACATTTGAGGAGCGGTGGAGGACACTCACCTGATGTTATC	1754
Db	1451	CGACCATCCACACTTCCATACCGAAGGGCGGCGCGGCCCATGCGCCGGACATCATC	1510
QY	1755	ACCATGGCAGCGAGCTCAATATTTCTACCTCTCTCCACACCCGCCACTATTTCCCTATACC	1814
Db	1511	ACCGCTCGGCCACCGAACAATTTGCGCTGCTGCAACCAACCAACGCTGCCCTACACC	1570
QY	1815	ATTAAATACGGTTGCGAACAATTTAGACATGCTCATGATGCCACCACTAGACAAACGC	1874
Db	1571	CTCAACACCATCGATGAACATCTCGATATGCTGATGCTGCGCACCAACCAACGCTGCCCGGAC	1630
QY	1875	ATCGCGAGGATTTACNAATTTTCTCAAAGCGGTATCGGCCCGGCTCTATCGCGCTGAA	1934
Db	1631	ATCGCGAGGAGCTGGGCTTTTGGCGAGTTCGCGCAATTCGCGGGAACCATCGCTGCGGAA	1690
QY	1935	GATGTGCTCCATGATATGGGTGTGATCGCGATGCAAGCTCGGATTCGCAAGCAATGGG	1994
Db	1691	GAGTGTGACGATCTCGGCGCTTCTCGCTCACTCTCCGATTCGAGGCCATGCGC	1750
QY	1995	CGTGACGGGAGTGTTCCTCGAACTTGGCAGACTGGGATAGAATAAAGAAATTT	2054
Db	1751	CGCGTGGGGAAGTATTTCTCCGACCTGGCAGGTGGCGCATCGCATGAAGTGCAGCGC	1810
QY	2055	GGTAAGCTTCTGAAAGTGGCAAGATAAACGATTAATTTCCGATTAAGCGCTACATCTCC	2114
Db	1811	GGAGCGTGGCGGAGGAGACCGGGGATAACGACAACTTCCGCTGAAGCGCTACATCGCC	1870
QY	2115	AAATACATATCAACCCCGCTTGACCCAGCGCTGAGCGAGTATATCGGCTCTGTGAA	2174
Db	1871	AAATACACCATCAACCGGCGCTGACCCAGCGCATCGCACAGAGTCGGATCCATTGAG	1930
QY	2175	GAGGCAAGATCGCCGACTTGGTGTGGAATCTCTGCTTTTGGCGTAAAAACCCAAA	2234
Db	1931	GTGGTAAAGCTGGCTGACTCGTGTCTGGTCAACAGCTTCTTCGGCGTGAACCGGCC	1990
QY	2235	ATCGTATCAAGCGCGTATGGTGTCTTCTGAAATGGCGGATTTCAAACGCTCTGTG	2294
Db	1991	ACCGTATCAAGCGCGCATGATCGCCATCGCGCGATGGGCGGATATCAATGCTCTATT	2050
QY	2295	CCACTCCCCAAACGGGTTTATACCGCAATGTTGGGCATCAGGCAAGCGGAAATTT	2354
Db	2051	CCGACCCCGAGCGGTGCACTACCGCCGATTTTGGCGGCTGGGCGCGCCGCAT	2110
QY	2355	GACACGAGTCACTTTTGTTCAAAGTCCCTATGAAAAATGGCGTGAAGAAAAAGCTG	2414
Db	2111	CATGCGCTCACTTCTGTCGAGCGCGGCGGCGCATGCGGTTGCGGAGCGGCTG	2170
QY	2415	GGCTTAGAGCGCCAAAGTTCTACCGGTCAAAAATCTCCGTAAACATCACCAAGAAAGCTTC	2474
Db	2171	AACCTGGCAGCGCATCGCGCTGTGAAGGCTGCGGTACCGTGCAGAAAGCCGACATG	2230
QY	2475	AAGTTCAACGACAAAAACGGCAAAATCACCGTCAATCGGAAAAACCTTCAGAGTCTTTGTA	2534
Db	2231	GTGCACAAACAGTCTGCGAGCCCTAAATCACCCGTCGACGCGCCAGACCTATGAGGTGCGGGTG	2290
QY	2535	GATGGCAAACTCTGCACTCTTAACCCACCTCGCAAGTGGCTCTAGCCCCAGCGCTACACT	2594

Db	2291	GATCGCAACTTATCACCAGCGAGCGGACAGAGCTTCTGCCGATGCGCAACGATATTTT	2350
Qy	2595	TTCTCTAGGCACAATGCCCTTT	2619
Db	2351	CTGTTTAAAGGAGCGGATGCTTT	2375
RESULT 11			
US-09-557-884-1/c			
: Sequence 1, Application US/09557884			
: Patent No. 6506581			
: GENERAL INFORMATION:			
: APPLICANT: Fleischmann et al.			
: TITLE OF INVENTION: The Nucleotide sequence of			
: the Haemophilus influenzae Rd Genome, Fragments			
: Thereof, and Uses Thereof			
: NUMBER OF SEQUENCES: 1			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Human Genome Sciences, Inc.			
: STREET: 9410 Key West Avenue			
: CITY: Rockville			
: STATE: MD			
: COUNTRY: USA			
: ZIP: 20850			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: 3 1/2 inch diskette			
: COMPUTER: Dell Pentium			
: OPERATING SYSTEM: MS DOS v6.22			
: SOFTWARE: ASCII Text			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/09/557,884			
: FILING DATE: 25-Apr-2000			
: CLASSIFICATION: <Unknown>			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: 08/476,102			
: FILING DATE: JUN-5-1995			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Michelle S. Marks			
: REGISTRATION NUMBER: 41,971			
: REFERENCE/DOCKET NUMBER: PB186P3			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: 301-309-8504			
: TELEFAX: 301-309-8439			
: INFORMATION FOR SEQ ID NO: 1:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 1830121 base pairs			
: TYPE: nucleic acid			
: STRANDEDNESS: double			
: TOPOLOGY: linear			
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:			
US-09-557-884-1			
Query Match 23.0%; Score 664.4; DB 3; Length 1830121;			
Best Local Similarity 57.1%; Pred. No. 1.le-159;			
Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9;			
Qy	183	AAATTTAACACAGAGTAAATAGGTGAACCTCACACCCAAAGACGAGAAAAAGTCTTGT	242
Db	564897	ATAATAAAATTAAGGAATGAAAATATGCATTAACCTTCAGAGAACCAAGAAAACTGATGC	564838
Qy	243	TATATTATCGGGCGAAGTGCTAGAAAGCCGAAGCAGAGGGCTTAAAGCTCAACCAAC	302
Db	564837	TTTTTCCTCGCGGGCGAACTTCGGCAAAACGCAAGCAGCGCGCTTAAATTTAAACTATC	564778
Qy	303	CCGAAGCATTGCTTACATTAGTGCCCATATATTGACGAAGCGCGCGTGGAAAAAAA	362
Db	564777	CAGAACTATTGCTTATATTCTAGTCAITTTACAGAGGCAGC---	564721
Qy	363	CCGTTGCCAGCTTATGGAAGAGTGCATGCACCTTTTGAAAAAAGATGAAGTAATGCCCG	422
Db	564720	GTGTAGCGGAAGTCATGCAATATATGCGCAACACACTTTTAAACCGTTGATGATGTCATGGAAG	564661

Db	562507	AACCACTGGATAGCGTACCATTTGGTTCAGCATATTTCTTCTTA	562462
RESULT 12			
US-09-643-990A-1/c			
; Sequence 1, Application US/09643990A			
; Patent No. 6528289			
; GENERAL INFORMATION:			
; APPLICANT: Robert D. Fleischmann			
; Mark D. Adams			
; Owen White			
; Hamilton O. Smith			
; J. Craig Venter			
; TITLE OF INVENTION: The Nucleotide sequence of			
; the Haemophilus influenzae Rd Genome, Fragments			
; Thereof, and Uses Thereof			
; NUMBER OF SEQUENCES: 1			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville,			
; STATE: MD			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3 1/2 inch diskette			
; COMPUTER: Dell Pentium			
; OPERATING SYSTEM: MS DOS v6.22			
; SOFTWARE: ASCII Text			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/643,990A			
; FILING DATE: 23-Aug-2000			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/487,429			
; FILING DATE: 1995-06-07			
; APPLICATION NUMBER: 08/426,787			
; FILING DATE: 1995-04-21			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Kenley K. Hoover			
; REGISTRATION NUMBER: 40,302			
; REFERENCE/DOCKET NUMBER: PH186P1C1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 301-610-5790			
; TELEFAX: 310-309-8439			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1830121 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:			
US-09-643-990A-1			
Query Match 23.0%; Score 664.4; DB 3; Length 1830121;			
Best Local Similarity 57.1%; Pred. No. 1.1e-159;			
Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9;			
Qy	183	AAATTTAACACAGGAGTATAGTGTAAACTCACACCAAGAGCAGAAAAGTCTTGT	242
Db	564897	ATATAAATTAAGGAATGAAATATGCATTTCAAGAGGAGCAGC	564838
Qy	243	TATATTATCGGGGAGTGGCTAGAAAGCCAAAGCAGAGGGCTTAAAGCTCAACCAAC	302
Db	564837	TTTTCTCGCGGGCGAACTTGGCGCAAAACCGAAAGCAGCGCGGTAAATTTAACTATC	564778
Qy	303	CCGAAGCCATTGCTTACATTAGTCCCATATTTATGGACGAAGCGCGCTGGAAGAAAAA	362
Db	564777	CAGAACTATTGCTTATATTCGTAGTCAATTTTCAAGAGGAGCAGC	564720
Qy	363	CCGTTGCCAGCTTATGGAAGTGCATGCATTTTGTGAAAAAGATGAAGTATGCCG	422
Db	564720	GTGTAGCGGAGTCAATGCAATATGCGCAACATTTTAAACGTTGATGATGATGCGAAG	564661

QY	423	GGGTGGGTAAATATGGTTCCCGATCTTAGTGTAGAAGCCACCTTTCTGTAGTGGTACGAAC	482	QY	1477	TTTTGGCAAAAGCAATAGCTCTAGCAAAAAACA	1536
Db	564660	GTGTGGCGGAAAATGGTTTCATGAAGTCCAGATTGAAGCTACTTTCCCGATGGCAGCAAC	564601	Db	563581	TTTTTGGTAAGGCAACCTGTTCAACCCCTAGAT	563522
QY	483	TTTGTAAGTGAATTTGGCCATC-----GAACCGATGAGCACTTCAAGCGGGGGAAGTG	538	QY	1537	CGATTGGTTTTAAATTTGCATGAAGACTTGGGGCACAACACCAAGTGCATCGATCACTGCT	1596
Db	564600	TTGTTACCGTGCATAATCCCAATCAGATAACCGWAGGTTGGCTTTAGCCCCACAAAATAAA	564541	Db	563521	CATTAGGTTTAAAAATCCAGAAAGACTGGGGTGCAACGCTCGCGTGAATTGATTCTGCCT	563462
QY	539	AAATTTGGTTGCGATAAAGACATCGAGCTCAATGCAGGCAAAAGAAAGTAACCGAACTTGAG	598	QY	1597	TGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACACCGATACAGTCAATGAGG	1656
Db	564540	AATATCAATGGTGGCTTAAGCCCAACCTACAGGAACAAGATATGATCCAGGCGCAAT	564481	Db	563461	TAAAGTAGCAGATGAATGGATATTCAAGTGGCCATTACACACAGACACGCTAAATGAAA	563402
QY	599	GTTTACTAATGAAGGGCTTAATCCTTGC--ATGTGGGTAGCCATTTCCACTTCTTTGAAG	656	QY	1657	CAGGTTATGTAGATGACACCCCTAAATGCATGAACGGGGCGGCATCCATGCTTACCACA	1716
Db	564480	ACCAATTAGCGGAAGCGGATTTCTCGCTAATGTGCGCAGAAAACCGTAAAAATCGAAG	564421	Db	563401	GTGGCTTTTGGAAAGACAGATGAAGCGGATTGATGACAGGATCAATCTACTATTTCATA	563342
QY	657	CTAACAGGCACATAAATTCGATCGTGAAAAGCCTATGGCAAAAGCCT-----AGATA	710	QY	1717	TTGAGGAGCGGTGGAGGACACTCACCTGATGTTATCACCATTCACCATGGCAGCGAGCTCAATA	1776
Db	564420	TAACAAAATTCAGCGGACCGCCCAATTCAGTTGGCTCGCATTTACCATTTTTTTGAAACCA	564361	Db	563341	CGGAGGCGCAGGTGGTGGTCATGCACCTGACATCATTTAAAGCAGCGATGTATTCAACG	563282
QY	711	TTCCCTCTGGCAACGCTACGCAATTTGGGCGAGCAAAACCGCAAAAGTGCAGTTGATTC	770	QY	1777	TTCTACCTCTCTCCACCAACCCGACTCGTCCGTTTACCBAAAACACCATTTGATGAACATT	1836
Db	564360	ATAATGCCCTTAAATTTGACCGCACTTTGGCACGTGGAATGCGCTTAAATGTTCCATCTG	564301	Db	563281	TATTACCTGCTTCAACCAACCGACTCGTCCGTTTACCBAAAACACCATTTGATGAACATT	563222
QY	771	CTCTTGTGGCAGTAAAAAAGTGAATGGCATGAACGGGCTTGTGAATAAACATCGCGGATG	830	QY	1837	TAGACATGCTCATGACATGCAACCTAGACAAAACGCAATCCGAGAGATTTTACAATTTT	1896
Db	564300	GCAATGGGTGCGTTTTTGAACCGGTGAAGTGAATCAGTGGAAATTAAGTTAGTTGCTTTGGTG	564241	Db	563221	TGGATATGTTGATGGTTTGCATCACTTAGATAAACCGCTGCCGGAAGACGTAGCTTTTG	563162
QY	831	AACGCCATAAACATAAAGCGCTTGAACAGCGGAAATCTCACGGATTTATCAAGTAAG--	888	QY	1897	CTCAAGCGGTATCGCCCCCGCTTATCGCGCTGAAGATGTCTCCATGATATGGGTG	1956
Db	564240	GTAACCA-AATCATTTATGTTTCCATAATCAAAATGATGGCAAAATTAAGGTAGGGCA	564182	Db	563161	CCGATAGCCGTATCCGCCCTGAAACCAATTGCAGCAGAAAGATATTTTGCATGATATGGCG	563102
QY	889	AGACTCCCATGAANAATGAANAACAAGAAATATGTAATATCTACGACCCACCAAGGCG	948	QY	1957	TGATCGGATGACAAGCTCGGATTCGCAAGCAATGCGGCGTGCAGGCGAAGTGAATTCCTC	2016
Db	564181	AGATGGCATTTAAACAATTTCAAGAGCGCAATATGTAGCAACTTATGCTCCAACAGTTTGGCG	564122	Db	563101	TCCTCTCCATTATGAGTTAGACTCTCAAGCGATGGAGCTATTGGCGAAGTCTGTTATTC	563042
QY	949	ATAAAGTGGCTTAGAGATACCGATCTTTGGGCGAGAAGTAGAACATGACTATATACCACT	1008	QY	2017	GAACTTGGCAGACTCGGGATGAAGATAAANAAGAAATTTGGTAAGCTTCTCGAAGATGGCA	2076
Db	564121	ATANAAGTCGTTTAGCGGATACCAATTTATGGGCAACCATTTGAACAAGATTTATGACCA	564062	Db	563041	GTACATGGCAAACTGCAGATAAGATAAGATAAATGCAACGTTGCTGAGCTAGGTAATGAAGGA	562983
QY	1009	ATGGCAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTAGGGGTATGGGTACAGAC-	1067	QY	2077	AAGATAACGATAATTTCCGCAATTAGCGCTACATCTCCAAATACACTATCAACCCCGCTT	2136
Db	564061	AAGGTATGAGTGAATTTGGTGGCGGTAAAAAGCGTGGTATGCTGATGGTATGGCTCAAAAGC	564002	Db	562982	-----AACGATAACTTCCGTATTAAACGATATATCGCAAAATACACCATCAACCCAGCAA	562928
QY	1068	-----AATAGCCCTGATGAANAACACCTAGATTTAGTTCATCACTTAACGGCATGATTA	1119	QY	2137	TGACCCACGCGGTGAGCGAGTATATCGGCTCTGTGGAAGAGGGCAAGATCGCCGACTTGG	2196
Db	564001	GTACGGCAACTCGCGCAATCCGAATGTATTGGATTTTGTGATTTACCAACGTGATGATCA	563942	Db	562927	TTGCACATGGTATTGCGGAGCATATTGGCTCGTTAGAGTGGGTAAAATTCGAGATATCG	562868
QY	1120	TCGA---CTACACCGGATTTACAAAGCCGACATTTGGGATTAANAACGCAAAATCCATG	1176	QY	2197	TGGTGTGGAATCTCTGCCCTTTTTTTGGCGTAAAAACCCCAAAATCGTGATCAAAAGCGGTATGG	2256
Db	563941	TTGATGCTAAATTAGGCAATTTACAAAGCCGATATTGGTATTCTGATTTGGCGGTATTGTGG	563882	Db	562867	TGTTATGGAACCGATGTTCTTTGGCGTAAACCTGNAAGTGGTGAATAAANAAGGCTTTA	562808
QY	1177	GCAATGGCAAGCAGGAAAACAAGACATGCAAGATGGCGTAAAGCCCTCATATGTCGTGG	1236	QY	2257	TGGTCTTCTCTGAATGGGCGATTTCTAACCGGTCTGTGCCCACTCTCCCAACCGGTTTATT	2316
Db	563881	GTATTGCAACAGCAGGTAAACCTTGACACCATGATGAACGTCAACCAAAATGATGATATCG	563822	Db	562807	TTAGCTATGCGAAAATGGGCGATCCAAATGCCTCAATTCCAACACCGCAACCTGTATTCT	562748
QY	1237	GTGTGGGCACAGAACACTAGCAGGGGAAGGTATGATTAATTAACGCTGGGGGAATCGATT	1296	QY	2317	ACCGGAAATGTTTGGGCACTACGCGAAAGGCGGAAAATTTTGACACCGAGCATCACTTTTGTGTT	2376
Db	563821	GTGCAAGCACGGAAGTTCATAACCGGTGCACATTTAATTTGCAACCGCTGGTGGTATCGATA	563762	Db	562747	ACCTCCAAATGTACGGTGCACAAGGCTTTAGCAACCGCAACCAACAGCAGTATTTCTTTGTTT	562688
QY	1297	CACACACCACTTCTCTTCTCCACAAACAATTCCTACCGCTCTAGCCAAATGGGCTTACAA	1356	QY	2377	CCAAAGTCGCTTATGAAATGGCGTGAAGAAAAGCTGGGCTTAGAGCGCCCAAGTTCTAC	2436
Db	563761	CCACATTCACTTATTTGTGTCACAAACAGCAACAACATGCAATTTGAAAGTGGCGTTACCA	563702	Db	562687	CACAAGCCGCTGAAAAGCTGATATTCTGTGCGAAGTTCGGTTTACAAAGAAACCATTTG	562628
QY	1357	CCATGTTTGGAGCGGCACAGGTCCTGTAGATGGCAGAAATGCGACTACTATCACTCCGG	1416	QY	2437	CGGTCAAAAACCTGCGGTAAACATCAACCAAGAAAGACTTCAAGTTTCAACGCAAAAACCGCAA	2496
Db	563701	CGTTAATTTGGTGGTGGAACTGGCCCTGCTGATGTCACACGCAACCACTTTGATCCCTG	563642	Db	562627	CTGTGAAAGGCTCGCCCAACGTAAGTAAAAAAGATCTGGTTCATAATGATGTAACACCAA	562568
QY	1417	GCAAAATGGAATTCGACCGCATTTGTCGCGCAGCAGGAAGATATTTCTAATGAAATGTCGGCT	1476	QY	2497	AAATCACCGTCGATCCGAAAACCTTCGAGGCTCTTTGTAGATGGCAAACTCTGCACCTCTA	2556
Db	563641	GCGCATGGTATATGGAACGATATGTTTCAAGCGGCAAGAGCCTTGGCCGGTAAACGTGGAT	563582	QY	2557	AACCCACCTCGCAAGTGCCTCTAGCCCGACGCTACACTTCTTCTTA	2602

/ LOCATION: (120038)..(120038)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (121344)..(121344)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (122167)..(122167)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (122336)..(122336)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (131340)..(131340)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (131360)..(131360)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (139910)..(139910)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (140398)..(140398)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (142750)..(142750)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (145058)..(145058)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (145171)..(145171)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (145942)..(145942)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (147197)..(147197)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (150841)..(150841)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (152500)..(152500)
/ OTHER INFORMATION: n equals a,t,c, or g
/ FEATURE:
/ NAME/KEY: misc_feature

Query Match 23.0%; Score 664.4; DB 3; Length 1830121;

Best Local Similarity 57.1%; Pred. No. 1.1e-159;

Matches 1397; Conservative 1; Mismatches 1012; Indels 36; Gaps 9;

QY 183 AAATTAAACACAGGAGTAGGTGAAACTCACACCCCAAGAGCAAGAAAGTTCTTGT 242

DB 564897 ATAATAAATTAGGAATGAAAATATGCATTCTTCCAGAGAACAGAAAACTGATGC 564898

QY 243 TATATTATCGGGCGGAAGTGGGTAGAAAAGCGCAAGCAGAGGGCTTTAAAGCTCAACCAAC 302

DB 564837 TTTTCTCTCGGGCGGAGTCTGGCGCAACAGCAACGCGGCTTAATTTAAACTATC 564778

QY 303 CCGAAGCCATTGCTTACATTAGTCCCATATTATGGACGAAGCGCGCTGGAAAAAAA 362
DB 564777 CAGAAACTATTGCTTATATTGCTAGTCAATTTTACAAGAGGCAGC---AAGAGAAGGAATGA 564721
QY 363 CCGTTGCCAGCTTATGGAAGAGTGCATGCACATTTTGAATAAAGATCAAGTAAGTAAATGCGG 422
DB 564720 GTGTAGCGGAAGTCATGCAATATGCGCAACACATTTTAAACCGTTGATGATGTCTATGGAAG 564661
QY 423 GGGTGGGTAATATATGTTTCCCGATCTAGGTGTAGAACCCACCTTTCTCTGATGTTAGCAAC 482
DB 564660 GTGTGGCGGAATGTTTCATGAAGTCCAGATTGAAGCTACTTTCCCGATGSCACGAAAC 564601
QY 483 TTGTAAGTGTAAATGCGCCCATC---GAACCAAGATGACATCTTCAAAAGCGGCGCAAGTG 538
DB 564600 TTGTTACCGTGCATAATCCAATCAGATAACCGWAGGGTGGGCTTTAGCCCAAAAATAAA 564541
QY 539 AAATTTGGTTGCGATAAAGACATCAGCTCAATCAGGCAAGAGATCAACCGAACTTGAG 598
DB 564540 AATATCAATGGTGGCTTAAAGCCACCTTACAGGAACAAGATATGATCCCGGCGGAAT 564481
QY 599 GTTACTAATGAAGGGCCTAAATCCTTGC--ATGTGGGTAGGCATTTCCACTTTCTTTGAAG 656
DB 564480 ACCAATTAGCCGAAGCGATATTTCTGCTAATGTGCGCAGAAAACCGTAAAAATCGAAG 564421
QY 657 CTAAACAGGCATAAAAATTCGATCGTGAATAAGCCTATGCAAAACGCCT-----AGATA 710
DB 564420 TAACAAATTCAGGCGACCGCCCAATTCAGTTGGCTCGCAATTACCATTTTTTTGAAACCA 564361
QY 711 TTCCCTCTGGCAACACGCTAGCGATTTGGGCGAGACAAACCCGCAAAAGTGAGTTGATTC 770
DB 564360 ATAAATGCCCTTAAATTTGACCGGCATTTTGGCAGTCGGAATGCGCTTAAATGTTCCATCTG 564301
QY 771 CTCTTGGTGGCAGTAAAAAAGTGTGCGCATGAACGGCTTGTGAAATAACATCGCGGATG 830
DB 564300 GCAATGCGGTGCGTTTTTGAACCCCGTGAAGTGAATCAGTGAATAGTTAGTTGCTTTTGGTG 564241
QY 831 AAGCCATAAACAATAAGCGCTTGACAAGGCGAAATCTCAGCGATTTATCAAGTAAGG-- 888
DB 564240 GTAACCA-NATCAATTTATGTTTCCATAATCAAAATGATGCAAAATTTAAGGTAGGGCA 564182
QY 889 AGACTCCCATGAAATGAAAAAACAAGAAATATGTAATACTACGGAACCCACCAAGGGG 948
DB 564181 AGATGGCATTAACAATTTCAAGAGCGCAATATGTAGCAACTTATGCTTCAACAGTTGGCG 564122
QY 949 ATAAAGTGGCTTAGGAGATACCGATCTTTGGGCGAAGTAGAATACATCACTATACCCT 1008
DB 564121 ATAAAGTCCGTTTTAGCGGATACCAATTTATGGGCAACCAATTTGAACAAGATTTATTGACCA 564062
QY 1009 ATGCGGAAGAACTTAAATTTGGCGCGGTAAAACTATCCGTGAGGATGGTTCAGAGC- 1067
DB 564061 AAGGTGATGATGTAAATTTGGTGGCGGTAAAGCGTGGTGATGGTATGCTCAAGCG 564002
QY 1068 -----AATAGCCCTGATGAAAAACCCCTAGATTTTAGTCATCATCAACCGGATGATTA 1119
DB 564001 GTACGGCAACTCGCGCAATCCGAATGTATTGGATTTGTGATTACCAACGATGATCA 563942
QY 1120 TCGA---CTACACGGGATTTAAACCGGCATTTGGATTTAAACCGCAAAATCCATG 1176
DB 563941 TTGATGCTAAATTTAGGCAATTTCAAAGCCGATATTGGTATTTCGTGATGGGCGTATTGTGG 563882
QY 1177 GCATTGGCAAGGCAAGGAAACAAGACATGCAAGATGCGTAAGCCCTCATATGGTCGTGG 1236
DB 563881 GTATTGGCAAGCAGGTAAACCTTGACCATGTGAATAGTCAACCAATATGATTTATCG 563822
QY 1237 GTGTGGCAGACAGACACTAGCAGGGGAAGTATGATTTATTACCGCTGGGGGAATTCGATT 1296
DB 563821 GTGCAAGCAGGAAGTTCTAAACGCTGCACATTTAATTTGCAACCGCTGCTGATCGATA 563762
QY 1297 CACACACCCACTCTCTTTTCCACAAACAATTTCCACCGCTCTAGCCAAATGGCGTTACAA 1356
DB 563761 CCCACATTCATTTTATTGTGCCAACCAAGCAACATGCAATTTGAAAGTGGCGGTACCA 563702
QY 1357 CCATGTTTGGAGGGGCACAGGTCTCTGTAGATGSCAAGATGCGCACTACTATCACTCCGG 1416

Db	562627	CTGTGAAAGGCTGCCGCAACGTTAGGTAAAAAAGATCTGGTTCATAATGATGATTAACACCAA	562568
QY	2497	AAATCACCGTCGATCCGAAAACCTTCGAGGTCTTTGTAGATGGCAAACTCTGCACCTCTA	2556
Db	562567	ACATTACTGTGTGATGCTGAACGTTATGAAGTTCGAGTGGACGGAGGTTAATTACCTGTG	562508
QY	2557	AACCCACCTCGAAAGTGCCCTCTAGCCAGCGCTACACTTTCTTCTTA	2602
Db	562507	AACCACTGGATAGCGTACCACTTGGTCAGCGATATTTCTCTATTCTA	562462
RESULT 14			
US-08-956-171E-22/c			
; Sequence 22, Application US/08956171E			
; Patent No. 6593114			
; GENERAL INFORMATION:			
; APPLICANT: Charles Kunsch			
; Gil H. Choi			
; Patrick S. Dillon			
; Craig A. Rosen			
; Steven C. Barash			
; Michael R. Fannon			
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences			
NUMBER OF SEQUENCES: 5256			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Human Genome Sciences, Inc.			
STREET: 9410 Key West Avenue			
CITY: Rockville			
STATE: Maryland			
COUNTRY: USA			
ZIP: 20850			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage			
COMPUTER: HP Vectra 486/33			
OPERATING SYSTEM: MSDOS version 6.2			
SOFTWARE: ASCII Text			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/956,171E			
FILING DATE: 20-Oct-1997			
CLASSIFICATION: <Unknown>			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 60/009,861			
FILING DATE: January 5, 1996			
APPLICATION NUMBER: 08/781,986			
FILING DATE: January 3, 1997			
ATTORNEY/AGENT INFORMATION:			
NAME: Mark J. Hyman			
REGISTRATION NUMBER: 46,789			
REFERENCE/DOCKET NUMBER: PB248P1			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (240) 314-1224			
TELEFAX: (301) 309-8439			
INFORMATION FOR SEQ ID NO: 22:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 5966 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: linear			
SEQUENCE DESCRIPTION: SEQ ID NO: 22:			
US-08-956-171E-22			
Query Match 23.0%; Score 662.8; DB 3; Length 5966;			
Best Local Similarity 55.5%; Pred. No. 2.1e-160;			
Matches 1437; Conservative 1; Mismatches 1093; Indels 59; Gaps 6;			
QY	70	TRCTATTATTAATTTTTTAATAATTACTTATATCATATAATAATAATATTACTTATA	129
Db	4148	TGCTTCAATATCGGATATGCTAAAAATAATAATTTCTAATTACACATTCGAATTTGTATA	4089
QY	130	TTAAAAAGTTAAATAAAGTAACGAAATTAGGACTATATCCCATGCCCTTTAAATAATTA	189
Db	4088	TTACTTGTTTTATCAACAATTAGGATTAATAATAAATTTTGTAAATTTGTATCTAATATGA	4029

QY 190 ACACAAGGAGTAATAGGTGAAACTCACACCAGAGCAAGAAAAGTTCTTGTATATTA 249
Db 4028 AATTGGGGGTTTTAAATTTGCAATTTTACACAAAGAGAGCAAAATTAATAGTTAGT 3969
QY 250 TCGCGGCGAAGTGGCTAGAAAGCGCAAGCAGAGGGCTTAAAGCTCAACCAACCGGAAGC 309
Db 3968 GCGCGGCGAAGTTGCGAGCTCGTGAAGCAGCCTGGTTTGAACTAAATCATCTCGAGGC 3909
QY 310 CATTTGCTTACATTAAGTCCCATATTAATGACGAAGCGCGCTGGAAGAAAACCGTTGC 369
Db 3908 ATTAGCTTTAATCAGCGATGAATATTAGAAAGTGCACGCGATGG---TAAGACCGTTGC 3852
QY 370 CCAGCTTATGGAAGAGTGCATGCATTTTGAAGAAAAGATGAAGTAATGCCCCGGGTGGG 429
Db 3851 AGAGTTAATGAGTTATGGTAGACAAATTTCTAAACAAAGAGAGTGTCAATGATGTTGTCGA 3792
QY 430 TAATATGGTTCCCGATCTAGGTGTAGAGCCACTTTCTCTGATGGTACGAACTTGTAAAC 489
Db 3791 ACACATGATTTACAGATATCGAAATCGAGGCTACGTTTCCCCGATGGTACTAAGTTAATCAC 3732
QY 490 TGTGAATTTGGCCCATCGAACCATAGTAGCACTTCAAAGCGGCG- GAAGTGAATTTTGGTT 548
Db 3731 AGTACATCACCTTATTTGTTAAAGGAGGCGAGTCAATATGATACAGAGAGAAATTTATACAA 3672
QY 549 GCGATAAAGACATCGAGCTCAATGCAAGGCAAGAAAGTAAACCGAATTTGAGTTACTAATG 608
Db 3671 AAAGTACAGAGGTTGAAATTTAATAACCATCATCTGAAACAGTTATCGAAGTTGAAATA 3612
QY 609 AAGGGCTTAATCTTTCATGTGGGTAGCCATTTTCACTTCTTTGAAGCTTAAACAGGCAC 668
Db 3611 CAGGAGATCGACCAATTTCAAGTGGGCTCAATTTTCATTTTATGAAGCAAAATGCGAGCAT 3552
QY 669 TAAAAATTCATCTGTGAAGAAAGCTATGCGAAACCGCTAGATATTCCTCTGCGCAACGC 728
Db 3551 TAGATTTGAAACGTGAATGACATATGGAATAATTTAGATATTCAGCTGGAGCAGCTG 3492
QY 729 TAGCATTTGGGCGAGCAACAAACCGCAAGTGCAGTTGATTCCTCTTGGTGGCAGTAAAA 788
Db 3491 TTTCGATTTGAACTGGGGATAAAAAGAGTTCAATTAGTTGAATGCTGGCAACGTA 3432
QY 789 AAGTGAATGGCATGAACGGGCTTTGTGAA---TAAACATCGCGAGTGAACGCCATTAACATA 845
Db 3431 AAATTTTGTGTTTCGTGTTATGGTCAATGGTCTATCGATGAGTCACTGCTATCGCC 3372
QY 846 AAGCGCTTGACAGGCGAATCTCAAGGATTAATC----- 880
Db 3371 CAACATGAAAAATGATGAATGCAAGGTGATTCGAGATACGGTGTGTAAGAACGTGA 3312
QY 881 -----AAGTAAAGGAGACTCCCATGAAATGAAAGAAACAAAGAAATATGTAATAC 928
Db 3311 ATAAAAAAGGAGGAAAAAGATCATGAGCTTTAAATGACCCAAATCAATATAGGAGCTT 3252
QY 929 CTACGACCCCAACAAAGGCGATAAGTGGCTTAGGAGATACCGATCTTTTGGGCGAAGT 988
Db 3251 ATACGGTCAACACTGTTGGAGATTCATTTGTTAGGTGATACGAATCTATTTGCTCAAT 3192
QY 989 AGAACATGACTATACCACTTATGGCGAAGAACTTAAATTTGGCGCGGTTAAACCTATCCG 1048
Db 3191 AGAAAAAGACTATGCGGTTTATGGTGAAGAAAGTACTTTTGGTGGTGAATCTATTAG 3132
QY 1049 TGAGGCTATGGGTACAGCAATAGCCC-----TGATGAAACACCCCTAGATTTAGT 1099
Db 3131 AGACGGTATGGCGCAAAATCTCTGTGTAAACGCTGATGAGTGAACGTTGCGACACCTTGT 3072
QY 1100 CATCACTAACCGCATGATTTACCTACATACACCGGATTTTACAAAGCGCAATTTGGGATTA 1159
Db 3071 CATTTCTAATGCGGTTATTCATGATACGATAAGTGGTTTAAAGCTGATATAGCACTAA 3012
QY 1160 AAACGCGAAAAATCCATGGCATTTGGCAAGGCGAGAAACAAAGCATGCAAGATGCGGTAAG 1219
Db 3011 AAATGGTTATATTTTCGCCATAGGTAATGCGCGCAACCCAGATATAATGGATAATGT--- 2955
QY 1220 CCTCATATGGTCTGGGTGGGCGACAGNAGCACTAGCAGGGGAAGGTATGATTATTAC 1279

Db 2954 ---CGACATTATATAGTTCAACAACAGATATCATTCGCTGAAGGTAAATCGTCAAC 2898
QY 1280 CGCTGGGCGAATCGAATTCACACACCACTTCTCTTTCCCAACAATTTCCCTACCGCTCT 1339
Db 2897 TGCTGGTGGTATTCATCTACTCATGTTTATTTAATCTTGAACAAGCAGAGGTGCGATT 2838
QY 1340 AGCCAATGGGCTTACAACCATGTTTGGAGGCGGCACAGGTCTGTAGATGGCAGCAATGC 1399
Db 2837 AGAAAGTGGTATTACGACTCATATTTGGTGGTGGTACTGCTGCTTCTGAGAGGTCTTAAAGC 2778
QY 1400 GACTACTATCACTCCGCGCAAAATCGAACTTGCACCGCATGTTGCGCGAGCAGAGAGTA 1459
Db 2777 AACAACTGTAACTCCAGGTCCATGGCATATTATAGAAATGTTAGAAAGTCTGCCAAGTTT 2718
QY 1460 TTCTATCAATGTGGGCTTTTGGGCAAGGCAATAGCTCTAGCAAAAAACAATTTGTAGA 1519
Db 2717 ACCGATTAATGTCCGTTTTACAGGTAAAGCAAGCAACAATCCACTGCACTCATTTGA 2658
QY 1520 ACAAAGTGAAGCGGCGCGATTTGGTTTAAATTCATGAAGACTGGGCGCAACCAACGA 1579
Db 2657 ACAAATCAATGCCGAGCAATTTGGATTAAAGATACATGAAGACTGGGCTGCAACCAATC 2598
QY 1580 TGGCATCGATCACTGCTTTGAGCGTGGCAGATGAATACGATGTGCAAGTTTGTATCCACAC 1639
Db 2597 TGCTTTGAGTCATGCAATTAGATGTTGCTGATGAATTTGATGTTCAATTTGCAATTATC 2538
QY 1640 CGATACAGTCAATGAGCGAGGTTATGTAGATGACACCCCTAAATGCAATGAACGGCGCGC 1699
Db 2537 AGATACTTTAAATGAAGCAGGATTTATGGAAGACACAATGGCTGCTGTTAAAGACGCTGT 2478
QY 1700 CATCCATCGCTTACCACATTTGAGGAGGGGTGGAGGACACTCACCTGATGTTTATCACCAT 1759
Db 2477 ACTTCATATGACATCTAGCTGAGGTGCTGGTGGCGGTCATGGCCTGATTTAATTTAAATC 2418
QY 1760 GGCAGGCGAGCTCAATATTTCTACCTCTCCACCAACCCCACTATTTCCCTATACCAATTA 1819
Db 2417 CGCTGCAATTTTCAAATATTTTACCTTCATCTACAAATCAAACTTTTGCCTTATACATATA 2358
QY 1820 TACGGTTGCAAGACACTTAGACATGCTCATGACATGCCACCCTAGCAACCAACGATCCG 1879
Db 2357 TACTGTAGATGAACATTTAGATATGGTAATGATTACTCACCATTTTAAATTCGCGCTTATCC 2298
QY 1880 CGAGGATTTCAATTTTCTCAAAGCGGTATCCGCGCCCGCTCTATCGCGGCTGAAGATGT 1939
Db 2297 TGAAGATATCGCATTCGAGATTCACGTTATTCGTAAAGAAACGATTCGAGCAGAGATGT 2238
QY 1940 GCTCCATGATATGGGTGTGATCGCGATGCAAGCTCGGATTCGCAAGCAATGGGCGGTGC 1999
Db 2237 TCTGCAAGATATGGGTGTATTCAGTATGATTTAGTTCCGATTTCAACAGCAATGGGCGGTGT 2178
QY 2000 AGCGAAGTGTATCTCGAACTTGGCAGATCGCGATGAAGATTAAGAAAGATTTGGTAA 2059
Db 2177 AGGTGAAGTAAATTAACAGCAACATGGCAAGTAGACATCGCATGAAGAAACCAACGTTGCTC 2118
QY 2060 GCTTCTGAGAGTGGCAAGATACGATAATTTCCGCAATTAAGCGCTACATCTCCAAATA 2119
Db 2117 TTTAGATGGTGAATTTGAAACATAATGATAATTCGCAATCAACGTTATATCGCTAAATA 2058
QY 2120 CACTATCAACCCCGCTTTGACCCCGGCTGAGCGAGTATATCGGCTCTGTGGAAGAGGG 2179
Db 2057 TACATTTAACCCAGCAATTAACATGTTATTTCTGAAATATGTAGGATCTATCGAGCCGG 1998
QY 2180 CAAGATCCGCACTTGGTGGTGAATCTCGCCCTTTTTCGGGTGAAACCCCAAAATCGT 2239
Db 1997 CAAACTAGCTGACATTTGCTTATGGGACCCCAATTTTCTTTGGGGTTAAACCTGGAATAGT 1938
QY 2240 GATCAAGGCGGATATGGTGGTCTTCTCTGAATGGCGGATTTCTAACGCTGTGTCGCCAC 2299
Db 1937 TGTAAGGCGGATTAATTAATCTCGCGTAAATGGCGATCAATGGTGTCTATACCTTAC 1878
QY 2300 TCCCAACCGGTTTATTTACCGCAAAATGTTTGGGCAATCAGCGCAAGGCGGAAATTTGACAC 2359

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:00:02 ; Search time 14.5164 Seconds
(without alignments)
1497.963 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLYYAGEVA.....ADRRHKHAKDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	613.5	51.9	238	2 B71977	urease alpha chain
2	610.5	51.7	238	1 URKCAP	urease EC 3.5.1.5
3	566.5	48.0	237	1 S35290	urease EC 3.5.1.5
4	483	40.9	228	2 D75586	urease, beta/gamma
5	439.5	37.2	840	1 URJB	urease EC 3.5.1.5
6	419.5	35.5	838	2 A96599	probable urease F1
7	398.5	33.7	835	2 T37939	urease EC 3.5.1.5
8	395	33.4	227	2 T29055	urease EC 3.5.1.5
9	281.5	23.8	100	1 A36138	urease EC 3.5.1.5
10	278.5	23.6	100	2 H70564	probable ureA prot
11	277.5	23.5	100	2 D85603	probable urease st
12	277.5	23.5	100	2 B90794	urease gamma subun
13	276.5	23.4	100	2 S74889	urease EC 3.5.1.5
14	271	22.9	105	2 B69729	urease (gamma subu
15	270.5	22.9	100	2 A36950	urease EC 3.5.1.5
16	269.5	22.8	100	2 S47102	urease EC 3.5.1.5
17	267.5	22.7	100	2 A64076	urease EC 3.5.1.5
18	266	22.5	137	1 S38484	urease EC 3.5.1.5
19	265.5	22.5	100	2 C90027	urease gamma subun
20	264.5	22.4	100	2 B47090	urease EC 3.5.1.5
21	263	22.3	124	2 C69729	urease (beta subun
22	262.5	22.2	100	1 S08478	urease EC 3.5.1.5
23	261.5	22.1	100	1 B43719	urease EC 3.5.1.5
24	259.5	22.0	100	2 AD2264	urease gamma chain
25	258.5	21.9	100	2 E83037	urease gamma subun
26	257.5	21.8	100	2 S42502	urease EC 3.5.1.5
27	256.5	21.7	100	2 D83681	urease gamma subun
28	251.5	21.3	100	2 AE2872	urease gamma subun
29	251.5	21.3	107	2 F97448	urease gamma chain

ALIGNMENTS

RESULT 1

B71977

urease alpha chain - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004

C:Accession: B71977

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: B71977

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-238 <ARN>

A:Cross-references: UNIPROT:Q9ZM24; UNIPARC:UPI0000137D6F; GB:AE001439; NIT

A:Experimental source: strain J99

C:Genetics:

A:Gene: ureA

C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K c

F:1-101/Domain: urease 11K chain homology <U11>

F:108-204/Domain: urease 12K chain homology <U12>

Query Match 51.9%; Score 613.5; DB 2; Length 238;

Best Local Similarity 53.1%; Pred. No. 1.3e-43;

Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKNOPEAIAIVISAHIMDEARRGKKTVAOLMER 60

Db 1 MKLTPKEDLKLMLHYAGELARRKKEGKIKLNTVEAVALISAHIMEARRGKKTVAOLMER 60

Qy 61 CMHFLKKDEVNPGVGNMVPDLGVEATFPDGTGLVTNVNPIEPDEHFKAGEVFGCKDIE 120

Db 61 GRTLLKPDVWDGVSAMHEVGIEAMFDGTGLVTNVNPIEPDEHFKAGEVFGCKDIE 119

Qy 121 LNAGKEVTELEVTNEGPKSLHGVSHFFFEANKALKFKREKAYKRLDIPSGNTLRIGAG 180

Db 120 INEKKAVSVKVKVGNDRPVQIGSHFFHFEVNRCLDFDREKTFGKRLDIASGTAVRFEFG 179

Qy 181 QTRKVLQIPGGSKKVIKMGVLNNIADRRHKHAKDKAKSHCF 224

Db 180 EKSVELIDIGNRRIRFGNALVDRQADNESKIALHRAKERGF 223

RESULT 2

URKCAP

urease (EC 3.5.1.5) 26K chain - Helicobacter pylori (strains 26695 and others)

N:Alternate names: urea amidohydrolase; urease alpha chain

C:Species: Helicobacter pylori

C:Date: 30-Sep-1991 #sequence_revision 15-Aug-1997 #text_change 09-Jul-2004

C:Accession: A38537; A64529; S07884; B55306; A41502; A61371; E49215

F;1-101/Domain: urease 11K chain homology <U1>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	51.7%;	Score	610.5;	DB 1;	Length	238;		
Best Local Similarity	52.7%;	Pred.	No. 2.4e-43;					
Matches	118;	Conservative	43;	Mismatches	62;	Indels	1; Gaps	1;

QY	1	VKLTPKQEKEFLLYAYAGEVARKKAEGKLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE	60
DB	1	MKLTTPKELDKLMLHYAGELAKRKEKGKILNLYVEAVALISAHIMEEARAGKKTAAELMQE	60

QY	61	CMHFLKDXEYMPGVGNVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAGEVKGCDDIE	120
DB	61	GRTWLKKENYMDGVASMIHEVGIEANFPDGTKLVTHTPTIEANGKLVPGSEL-FLKNEDIT	119

QY	121	LNAGKEVTELEVTNEGPKSLHVSGSHFFPEANKALFKDREKAYGKKLDTPSGNLTIRIGAG	180
DB	120	INEGKAVSVKVNVDGRPVQIOGSHFFFEVNRCLDFDREKTKGKKLDIASGTAVRPEFG	179

QY	181	QTRKVQIPLIGGSKKIVGMNGLVNNIADERRHKHKALDKAKSHGF	224
DB	180	EKSVELIDIGGNRRIFGPNALVDROADNESKIALHRAKERGF	223

RESULT 3

S35290
urease (EC 3.5.1.5) 26K chain - Helicobacter felis
N:Alternate names: urease alpha chain
C:Species: Helicobacter felis
C:Date: 31-Dec-1993 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S35290
R:Ferrero, R.L.; Labigne, A.
Mol. Microbiol. 9, 323-333, 1993
A:Title: Cloning, expression and sequencing of Helicobacter felis urease genes.
A:Reference number: S35290; MUID:94018627; PMID:8412663
A:Accession: S35290
A:Molecule type: DNA
A:Residues: 1-237 <PER>
A:Cross-references: UNIPROT:Q08715; UNIPARC:UPI0000137D6C; EMBL:X69080; NID:G593
C:Genetics:
A:Gene: ureA
C:Complex: heterodimer
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide
C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease
C:Keywords: heterodimer; hydrolase; metalloprotein
F;1-101/Domain: urease 11K chain homology <U1>
F;108-204/Domain: urease 12K chain homology <U12>

Query Match	48.0%;	Score	566.5;	DB 1;	Length	237;		
Best Local Similarity	50.0%;	Pred.	No. 1.1e-39;					
Matches	112;	Conservative	43;	Mismatches	68;	Indels	1; Gaps	1;

QY	1	VKLTPKQEKEFLLYAYAGEVARKKAEGKLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE	60
DB	1	MKLTTPKELDKLMLHYAGELAEALARGLVNLNTEAVALISGRVMEKARDGNKSVADLMQE	60

QY	61	CMHFLKDXEYMPGVGNVPDLGVEATFPDGTKLVTNNWPIEPDEHFKAGEVKGCDDIE	120
DB	61	GRTWLKKENYMDGVASMIHEVGIEANFPDGTKLVTHTPTVEDNGKLAPGEV-FLKNEDIT	119

QY	121	LNAGKEVTELEVTNEGPKSLHVSGSHFFPEANKALFKDREKAYGKKLDTPSGNLTIRIGAG	180
DB	120	INAGKEAISLVKNVDGRPPVQVQSGSHFFFEVNRKLLDFDRAKSPCKRLDIASGTAVRPEFG	179

QY	181	QTRKVQIPLIGGSKKIVGMNGLVNNIADERRHKHKALDKAKSHGF	224
DB	180	EKSVELIDIGGNKRIYGFNSLVDROADGCKLGLKRAKEKGF	223

RESULT 4

D75586
urease, beta/gamma subunit - Deinococcus radiodurans (strain R1)

A;Reference number: JT0001
A;Accession: JT0001
A:Molecule type: protein
A;Residues: 1-246,'R',248-257,'P',259-268,'S',270-840 <TA2>
A;Cross-references: UNIPARC:UPI0000172D9F
R;Mamiya, G.; Takishima, K.; Masakuni, M.; Kayumi, T.; Ogawa, K.
J. Protein Chem. 6, 55-59, 1987
A>Title: Complete amino acid sequence of jack bean urease.
A;Reference number: A60883
A;Accession: A60883
A:Molecule type: protein
A;Residues: 1-246,'R',248-840 <TA3>
A;Cross-references: UNIPARC:UPI0000172D9E
R;Takishima, K.; Mamiya, G.
Protein Seq. Data Anal. 1, 103-106, 1987
A>Title: Location of the essential cysteine residue of jack bean urease.
A;Reference number: S03364; PMID:88190054; PMID:3447159
A;Accession: S03364
A:Molecule type: protein
A;Residues: 591-637 <TA4>
A;Cross-references: UNIPARC:UPI0000172DA0
C;Comment: Each chain of the hexamer binds two nickel ions.
C;Superfamily: urease; urease 1LK chain homology; urease 12K chain homology; urease 12K chain homology; urease 62K chain homology; hydrolase; metalloprotein; nickel
C;Keywords: homohexamer; urease 1LK chain homology <Uil>
F;1-101/Domain: urease 1LK chain homology <Uil>
F;134-230/Domain: urease 12K chain homology <Uil2>
F;274-823/Domain: urease 62K chain homology <U62>
F;407,409,490,633/Binding site: nickel 2 (His, Lys, Asp) #status predicted
F;490,519,545/Binding site: nickel 1 (Lys, His, His) #status predicted
F;490/Binding site: carbon dioxide (lys) (covalent) (by urease activase) #status predicted

Query Match 37.2%; Score 439.5; DB 1; Length 840;
Best Local Similarity 39.6%; Pred. No. 1.7e-28;
Matches 99; Conservative 47; Mismatches 77; Indels 27; Gaps 3;

QY 1 VKLTPEQEKFLLYAGVEAVARKKAEGKLKNQEPFAIYAIVISAHINDEARRGKTVAQLME 60
:|::||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 1 MKLSPREVEKGLHNGAYLAQRKLARGVRLNYTEAVALLASQIMEYARDGEKTVQMLCL 60

QY 61 CMHFELKKDVMPGVGNMVDLGVEATFPDGCTKLVTVMNPIDPDHFHKAGEVKFG----- 114
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 61 GOHLGRROVLPAVPHLLNAVQEATFPDGTKLVTVDHPIS-RENGELQEAALFGSLLPVP 119

QY 115 -----C-DKDIELNAGKEVTELEVNKGPSLHVSHGFPEAKNA 154
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:|:|:|:|:|:|:|:|:|:
DB 120 SLDKFAETKEDNRIPCEILCEDDECLTINIGRKAVILKVTSKDRPIQVGSHVFIEVNFY 179

QY 155 LKFDREKAYGKRLDTPSGNTLRIGAGTRKVLPLGGSKVIGMNLVINNIADERHKHK 214
|||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 180 LTFDDRKYGMRLNTAAAGTAVREPDPGCKSVTLWSIEGNKVIRGNGAINADGPVNETNLEA 239

QY 215 ALDKAKSHGF 224
|:~::~:||
DB 240 AMHAVRSKGF 249

RESULT 6
A96699
probable urease F12B7.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C;Accession: A96699
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberrg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A96141; PMID:21016719; PMID:11130712

A:Accession: A96699
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-838 <STO>
A:Cross-references: UNIPARC:UPI00000ABAGF; GB:AE005173; NID:g6382524; PIDN:AAF07806.1; C
C:Genetics:
A:Gene: F12B7.10
A:Map position: 1
C:Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K

Query Match 35.5%; Score 419.5; DB 2; Length 838;
Best Local Similarity 37.2%; Pred. No. 7.7e-27;
Matches 93; Conservative 50; Mismatches 80; Indels 27; Gaps 3;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLLPREIEKLEHQAGFLAQKRLARGIRLNTAEVALIATQILEFIRGDKSVAEIMDI 60
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGKLVTVNWPTE----- 101
Db 61 GRQLLGRQVLPVLLHLLTVQVEGTFRDGTGKLVTVNWPTELESLGNLELALHGSFLPVPS 120
QY 102 ----PDEH---FKAGEVKGCDKDIELNAGKEVTELEVTNEGPKSLHVGSHFFPEANKA 154
Db 121 LDKFPEVHEGVIPGDMKYG-DGSIIRHGRKAWLKVNTGDRPVQVQSHYHFISVNP 179
QY 155 LKFDREKAYGKRLDIPSGNLTIRIGAGQTRKQVLIPLGGSKKVGIMGNLVNNTADERHKH 214
Db 180 LVFDRKALGMRLNIPAGTAVRPEPGRKSVLLVNIIGNKVIKGGNGIVDGLVDDVNWTV 239
QY 215 ALDKAKSHGF 224
Db 240 LMETWERRGF 249

RESULT 7
T37939
urease (EC 3.5.1.5) - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 12-Jul-2004
C:Accession: T37939
R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, August 1999
A:Reference number: 221755
A:Accession: T37939
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-835 <MCD>
A:Cross-references: UNIPROT:O00084; UNIPARC:UPI0000137DA9; EMBL:AL109820; PIDN:CAB52575.
A:Experimental source: strain 972h-; cosmid c1952
C:Genetics:
A:Gene: SPDB:SPAC1952.11c
A:Map position: 1
C:Superfamily: urease; urease 11K chain homology; urease 12K chain homology; urease 62K
C:Keywords: hydrolase

Query Match 33.7%; Score 398.5; DB 2; Length 835;
Best Local Similarity 36.1%; Pred. No. 4.2e-25;
Matches 91; Conservative 38; Mismatches 94; Indels 29; Gaps 2;

QY 3 LTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEECM 62
Db 1 MQPRELHKUTLHQLGSLAQKRLCRGKLNKLEATSLASQIQEYVRDGNHVSADLMSLGK 60
QY 63 HFLKDEVMPGVGNMVPDLGVEATFPDGTGKLVTVNWPTE-----EPDE 104
Db 61 DMLGKRHVQPNVHLLHEIMEATFPDGTGTVLIHDPICITDGNLHVALGSLFPTPSQE 120
QY 105 HFKAGEVKEGC-----DKDIELNAGKEVTELEVTNEGPKSLHVGSHFFPEANK 153
Db 121 LFPLEBKLYAPENSPGVFVLEGEIELLNPLRPTPIEVNRMGDRPIQVQSHYHFIEETNE 180
QY 154 ALKFDREKAYGKRLDIPSGNLTIRIGAGQTRKQVLIPLGGSKKVGIMGNLVNNTADERHKH 213

Db 181 KLCFDRSKAYGKRLDIPSGTAIRFEPGVGMKIVNLIPIGGAKLIQGGNSLKGVFDDSRTR 240
QY 214 KALDKAKSHGFI 225
Db 241 EIVDNLMKQGF 252

RESULT 8
T29055
urease (EC 3.5.1.5) beta/gamma chain - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 06-Jan-2003
C:Accession: T29055
R:Redenbach, M.; Kieser, H.M.; Denapaite, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopwood
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: 220556; MUID:97000351; PMID:8843436
A:Accession: T29055
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-227 <RED>
A:Cross-references: UNIPARC:UPI00000DACF4; EMBL:AL031124; PIDN:CAA19973.1
C:Genetics:
A:Gene: ureAB
C:Superfamily: urease, fused gamma/beta subunit; urease 11K chain homology; urease 12K chain
C:Keywords: hydrolase

Query Match 33.4%; Score 395; DB 2; Length 227;
Best Local Similarity 39.7%; Pred. No. 1.7e-25;
Matches 92; Conservative 45; Mismatches 79; Indels 16; Gaps 4;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MRLTPTERDRLLLFGRALARRARGRLNVPENTALIADTVCEAARDGAR-LQAQIER 59
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGKLVTVNWPTEPDESHFKAGEVKGFGCDKDI 120
Db 60 ARSVLGPDDVLPGVADVTVTEVHVEAVFDGSRSLAVADPV-----GGGGLGDADFGA 111
QY 121 LNAGKEVTE-----LEVTNEGPKSLHVGSHFFFEANKALKFDREKAYGKRLDIPSGNT 174
Db 112 LLPGHRPEPEAALRLPVTNTATVPVSVTSHFFFEANPRLDPRGRAYGMRLAVPAGSS 171
QY 175 LRIGAGQTRKQVLIPLGGSKKVGIMGNLVNNTAD-ERHKKALKDKAKSHGFI 225
Db 172 VRFPGERVEGLVPIGGARVAIGFAGLVDPGLDPAGAREALRAAACGYL 223

RESULT 9
A36138
urease (EC 3.5.1.5) 11K chain [validated] - Klebsiella pneumoniae
N:Alternate names: urea amidohydrolase; urease chain A; urease gamma chain
C:Species: Klebsiella pneumoniae
C:Date: 30-Nov-1990 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: A36138; S32938
R:Mulrooney, S.B.; Hausinger, R.P.
J. Bacteriol. 172, 5837-5843, 1990
A:Title: Sequence of the Klebsiella aerogenes urease genes and evidence for accessory pr
A:Reference number: A36138; MUID:91008957; PMID:2211515
A:Accession: A36138
A:Molecule type: DNA
A:Residues: 1-100 <MUL>
A:Cross-references: UNIPROT:P18316; UNIPARC:UPI00001103C0; GB:M36068; NID:g149335; PIDN:
R:Collins, C.M.; Gutman, D.M.; Laman, H.
Mol. Microbiol. 8, 187-198, 1993
A:Title: Identification of a nitrogen-regulated promoter controlling expression of Klebs
A:Reference number: S32937; MUID:93268094; PMID:8497192
A:Accession: S32938
A:Molecule type: DNA
A:Residues: 1-66, 'K', 68-100 <COL>
A:Cross-references: UNIPARC:UPI0000137D9A; EMBL:L07039; NID:g149330; PIDN:AAA25147.1; PI
R:Jabri, E.; Karplus, P.A.

submitted to the Brookhaven Protein Data Bank, June 1995

A:Reference number: A65977; PDB:1KRA
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 1-100
R:Jabri, E.; Carr, M.B.; Hausinger, R.P.; Karplus, P.A.
Science 268, 998-1004, 1995
A:Title: The crystal structure of urease from *Klebsiella aerogenes*.
A:Reference number: A56340; MUID:95273988; PMID:7754395
A:Contents: annotation; X-ray crystallography, 2.3 angstroms
C:Genetics:
A:Gene: ureA
C:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two
C:Superfamily: urease, gamma subunit; urease 11K chain homology
C:Keywords: heterotrimer; hydrolase; metalloprotein
F:1-100/Domain: urease 11K chain homology <U1>

Query Match 23.8%; Score 281.5; DB 1; Length 100;
Best Local Similarity 59.0%; Pred. No. 1.7e-16;
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59

Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTKLVTNNPI 100
Db 60 GRHVLREQWMEGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 10

H70664

probable ureA protein - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70664
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-100 <COL>
A:Cross-references: UNIPROT:P50043; UNIPARC:UPI0000137D9C; GB:Z83859; GB:AL123456; NID:G
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: ureA
C:Superfamily: urease, gamma subunit; urease 11K chain homology
F:1-100/Domain: urease 11K chain homology <U1>

Query Match 23.6%; Score 278.5; DB 2; Length 100;
Best Local Similarity 58.0%; Pred. No. 3e-16;
Matches 58; Conservative 18; Mismatches 23; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPHEQRLLSYAEALARRRRARGLNHPETAVTADHILEGARDG-RIVAEILMAS 59

Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTKLVTNNPI 100
Db 60 GREVLGRDDWMEGPEMLAEVQVEATFPDGTKLVTNNPI 99

RESULT 11

D85603

probable urease structural subunit A (gamma) [imported] - *Escherichia coli* (strain O157:
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85603; E85654

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Iim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <STO>
A:Cross-references: UNIPROT:Q8XAG2; UNIPARC:UPI00000D5DF7; GB:AE005174; NID:g12513945;
A:Experimental source: strain O157:H7, substrain EDL933
A:Accession: E85654
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <ST2>
A:Cross-references: UNIPARC:UPI00000D5DF7; GB:AE005174; NID:g12514462; PIDN:AAG55697.1;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: ureA; ureA_2
C:Superfamily: urease, gamma subunit; urease 11K chain homology

Query Match 23.5%; Score 277.5; DB 2; Length 100;
Best Local Similarity 57.0%; Pred. No. 3.6e-16;
Matches 57; Conservative 20; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVAALMEE 59

Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTKLVTNNPI 100
Db 60 GRHVLREQWMEGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 12

B90794

urease gamma subunit [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 050995
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90794
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90794
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <HAY>
A:Cross-references: UNIPROT:Q8XAG2; UNIPARC:UPI00000D5DF7; GB:BA000007; PIDN:BA034745.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs1322
C:Superfamily: urease, gamma subunit; urease 11K chain homology

Query Match 23.5%; Score 277.5; DB 2; Length 100;
Best Local Similarity 57.0%; Pred. No. 3.6e-16;
Matches 57; Conservative 20; Mismatches 22; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MELTPREKDKLLFTAAALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVAALMEE 59

Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTKLVTNNPI 100
Db 60 GRHVLREQWMEGPEMIPDIQVEATFPDGSKLVTNNPI 99

RESULT 13

S74889

urease (EC 3.5.1.5) 11K chain - *Synechocystis* sp. (strain PCC 6803)
N:Alternate names: protein slr1256; urease gamma chain
C:Species: *Synechocystis* sp.


```

Db      1  MKLTPVQEQKLLIFRAGELAKORKARGVLLNTPPEAAAYITCFIMGARDG-KGVAELMEA 5
Qy      61  CMHFLKKDEVMPGVGNWPDGLGVEATFPDGTGLVTVNWPIDEPDEHFKAQEVK 112
Db      60  GRHVLTEKDMGVGPEMLDSIQVEATFPDGVKLVTVHQPI-----SAEVK 104

RESULT 15
A36950
  urease (EC 3.5.1.5) 11k chain - Bacillus sp. (strain TB-90)
N;Alternate names: ureA protein
C;Species: Bacillus sp.
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C;Accession: A36950
R;Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T.
J. Bacteriol. 176, 432-442, 1994
A;Title: Cloning, sequencing, and expression of thermophilic Bacillus sp. str
A;Reference number: A36950; MUID:94117379; PMID:8288539
A;Accession: A36950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-100 <MAE>
A;Cross-references: UNIPROT:Q07399; UNIPARC:UPI0000137D95; GB:D14439; NID:G39
C;Superfamily: urease, gamma subunit; urease 11k chain homology
C;Keywords: hydrolase
F;1-100/Domain: urease 11k chain homology <U11>

      Query Match      22.9%; Score 270.5; DB 2; Length 100;
      Best Local Similarity 55.0%; Pred. No. 1.4e-15;
      Matches 55; Conservative 19; Mismatches 25; Indels 1; Gaps

Qy      1  VKLTPKEQKFLLYAGEVARKKRKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db      1  MKLTSREMEKLMIVVAADLARRRKERGLKLNYPEAVAMITYEVLEGARDG-KTVAQLMQY 59

Qy      61  CMHFLKKDEVMPGVGNWPDGLGVEATFPDGTGLVTVNWPPI 100
Db      60  GATILTKEDVMGVAEMIPDIQIEATFPDGTGLVTVHDPI 99

Search completed: November 28, 2005, 08:20:40
Job time : 15:5164 secs

```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	652	55.2	234	1	UREA_HELBI	Q8gh98 helicobacte
2	622	52.7	234	1	UREA_HELHE	P42822 helicobacte
3	613.5	51.9	238	1	UREA_HELPI	Q8zm24 helicobacte
4	611.5	51.8	238	2	Q8RNU5_HELPI	Q8rnus helicobacte
5	611.5	51.8	238	2	Q8JNU1_HELPI	Q8snj1 helicobacte
6	610.5	51.7	238	1	UREA_HELHE	P14916 helicobacte
7	610.5	51.7	238	1	UREA_HELHE	Q83ni3 helicobacte
8	610.5	51.7	238	2	Q83NI3_HELPI	Q83ni3 helicobacte
9	609.5	51.6	238	2	Q83NI8_HELPI	Q83ni8 helicobacte
10	608.5	51.5	238	2	Q833M9_HELPI	Q833m9 helicobacte
11	607.5	51.4	238	2	Q82S55_HELPI	Q832s5 helicobacte
12	607.5	51.4	238	2	Q83NI4_HELPI	Q83ni4 helicobacte
13	607.5	51.4	238	2	Q83NJ0_HELPI	Q83nj0 helicobacte
14	607.5	51.4	238	2	Q83NJ2_HELPI	Q83nj2 helicobacte
15	605.5	51.3	238	2	Q84F76_HELPI	Q84f76 helicobacte
16	604.5	51.2	238	2	Q83NI9_HELPI	Q83ni9 helicobacte
17	603.5	51.1	238	2	Q83NH7_HELPI	Q83nh7 helicobacte
18	603.5	51.1	238	2	Q83NI2_HELPI	Q83ni2 helicobacte
19	602.5	51.0	238	2	Q833P5_HELPI	Q833p5 helicobacte
20	601.5	50.9	238	2	Q83NI0_HELPI	Q83ni0 helicobacte
21	601.5	50.9	238	2	Q83NI1_HELPI	Q83ni1 helicobacte
22	599.5	50.8	238	2	Q83NH8_HELPI	Q83nh8 helicobacte
23	599.5	50.8	238	2	Q83NI6_HELPI	Q83ni6 helicobacte
24	598.5	50.7	238	2	Q83NI7_HELPI	Q83ni7 helicobacte
25	596.5	50.5	238	2	Q83NH9_HELPI	Q83nh9 helicobacte
26	566.5	48.0	237	1	UREA_HELPE	Q08715 helicobacte
27	544.5	46.1	213	1	UREA_HELWU	P50044 helicobacte
28	542.5	45.9	347	2	Q7PED9_ANOGA	Q7ped9 anopheles g
29	540.5	45.8	225	1	UREA_HELHP	Q83pj5 helicobacte
30	498	42.2	219	1	URE23_SULTO	Q972v9 sulfobact
31	491	41.6	223	2	Q5FPB21_CAMLA	O5fb24 campylobact

```

Query Match          55.2%; Score 652; DB 1; Length 234;
Best Local Similarity 57.3%; Pred.No.9.4e-4;
Matches 129; Conservative 39; Mismatches 55; Indels      2; Gaps      2;

QY    1 VKLTPEQEKFLLYYAGEVARKEKAEGLKLNQPEAIAYISAHIMDEARRGKTVAQLMEE 60
DB     :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
1 MKLTPLELDKMLHUYAGELAKKRANGVKUNYTEVALISAHVMEEARAKKSVDLMQE 60
DM     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY    61 CMHFLLKDDEVMPGVGNVDPDLGVEATFPDGTCLTVNWPPIEDSHFKA-GEVFEGCDKI 119
DB     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
61 GRTLRKADDWPGVAAHMEHVIGIEANFPDGTCLVTIHTPVEDGGHKLPAGEVLK-NEDI 119
DM     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY    120 ELNAGKEVTELVNTNGPKSLHGVSHPHFPAENKALKAFREKAYKKRLDTIPSGNTLRIGA 179
DB     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
120 TLNAGQAATTLEVHNKGDRPVQGSHFFHEFVNKLLEFDREKAYKKRLDIASGTAVRFE 179
DM     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
QY    180 GQRKVQLIPLGSKKVIGNMGVLNNIADSRHHKHKALDKAKSHGF 224
DB     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
180 GEKKTVELTIQGNQRIYGNSLVRQADTDGKKLALKRAKEGHF 224
DM     |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:

RESULT 2
UREA_HELHE STANDARD; PRT; 234 AA.
AC P42822;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase alpha subunit).
DE Names:urea;
OS Helicobacter heilmannii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=35817;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=2;
RX MEDLINE=94222523; PubMed=8168924;
RA Soinick J.V., O'Rourke J., Lee A., Tompkins L.S.;
RT "Molecular analysis of urease genes from a newly identified uncultured
RT species of Helicobacter.";
RL Infect. Immun. 62:1631-1638(1994).
CC -|- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -|- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic [by similarity].
CC -|- SIMILARITY: In the N-terminal section; belongs to the urease gamma
CC subunit family.
CC -|- SIMILARITY: In the C-terminal section; belongs to the urease beta
CC subunit family.
CC -|- CAUTION: In Helicobacter the alpha subunit is what is known, in
CC other bacteria as the beta subunit.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; L25079; AAA65722.1; -; Genomic_DNA.
DR HSPP; P14916; IE92.
DR SMR; P42822; 1-224.
DR HAMAP; MF_00739; fused; 1.
DR InterPro; IPR002015; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gammabeta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSFO01225; Urease_gammabeta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.

```

```
KW Hydrolase.
FT REGION      1    102       Urease gamma.
FT REGION      103   234       Urease beta.
SQ SEQUENCE    234 AA; 25746 MW; CD5889E3D789C7F8 CRC64;

Query Match          52.7%; Score 622; DB 1; Length 234;
Best Local Similarity 54.7%; Pred. No. 3.2e-44;
Matches 123; Conservative 41; Mismatches 59; Indels 2; Gaps 3;

QY 1 VKLTPEQEKFLLIYAGEVARKKAEGLKLNPQEAIAIYISAHIMDEARRGKTVAQLMEE 60
   :|||:::|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:
Db 1 MKLTPRELDKLMHLYAGELAKQRKAKGIKLYNTEAVALISAHVMEERARAGKSVDLMQE 60
   :|||:::|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:

QY 61 CMHFLLKKDEVMPGVGNVPDLPGVEATFPDGTKLVTNVNWPIEP-DEHPKAGEVKFGCDKI 119
   ||:|||::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 61 GRTLRKADDVMPGVAMHIEVGTEAGFPDGTKLVTIHTPVEAGSDKLAPGEVLK-NEDI 119
   ||:|||::|:|:~::~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:

QY 120 ELNAGKEVTELVNTEGGPSKSLHVSGSHFHFEPAEKALKAFORAKYGRRLDIPSGNTLRIGA 179
   |||||::|:|:~::~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
Db 120 TLNAGHAAVLQKVNGDRPVQVGSFHFFEVFNKLLDFDREKAYGRRLDIASGTAVRPEP 179
   |||||::|:|:~::~|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:

QY 180 GQTRKYQLIPLGGSKVIGMGVLVNNIADHRHKHALDKAKSHGF 224
   |:~::~|:|:|:|:|:~::~|:|:|:|:|:~::~|:|:|:|:|:~::~|:|:
Db 180 GESEKTVELIDIGNKRIYGFNALVDRQADHDGKKLKRAKEKH 224
   |:~::~|:|:|:|:|:~::~|:|:|:|:|:~::~|:|:|:|:|:~::~|:|:

RESULT 3
UREA_HELPJ STANDARD; PRT; 238 AA.
AC Q9ZMZ4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) [Urea amidohydrolase alpha subunit].
GN Name-urea; Synonyms=hpaA; OrderedLocusNames=jhp0068;
OS Helicobacter pylori J99 (Campylobacter pylori J99);
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI TaxID=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Urea + H(2O) = CO(2) + 2 NH(3).
CC -!- SUBUNIT: Heterodimer of an alpha subunit and a beta subunit (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to the urease gamma
CC subunit family.
CC -!- SIMILARITY: In the C-terminal section; belongs to the urease beta
CC subunit family.
CC -!- CAUTION: In Helicobacter the alpha subunit is what is known, in
CC other bacteria as the beta subunit.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
CC ENBL; AE001446; AAD05652.1; -; Genomic_DNA.
DR DR
DR PTR; B71977; B71977.
DR HSPSP; P14916; 1B9Z.
DR SMR; Q9ZMZ4; 1-238.
DR HAMAP; MF_00739; fused; 1.
DR InterPro; IPR002019; Urease beta.
DR DR
```

```
DR InterPro; IPR002026; Urease_gamma_reg.
DR InterPro; IPR008223; Urease_gamma_beta.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
KW Complete proteome; Hydrolase.
FT REGION 1 102 Urease gamma.
FT REGION 103 238 Urease beta.
SQ SEQUENCE 238 AA; 26568 MW; A10B9DC4156C0561 CRC64;

Query Match 51.9%; Score 613.5; DB 1; Length 238;
Best Local Similarity 53.1%; Pred. No. 1.7e-43;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEKGKILNYYEAVLISAHIMEEARAGKKTAAELMQE 60

Qy 61 CMHFLKDEVMGPNVVDLGVETATPDGKLVTVNWPIDPEHFKAGEVKFGCDKDI 120
Db 61 GRTLKPDVMDGVASMIHEVGIEAMPDPDGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFPEANKALFKDREKAYGKRLDIPSGNLTIRIGAG 180
Db 120 INEGKAVSVKVNVDGPRVQIGSHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179

Qy 181 QTRKVQLIPLGSKKVGIMGNLVNNIADERHGHKALDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223

RESULT 4
Q8RNU5 HELPY
ID Q8RNU5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q8RNU5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease A subunit.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=20123532; PubMed=10660136;
RA Opazo P., Muller I., Rollan A., Valenzuela P., Yudelevich A.,
RA Garcia-de la Guarda R., Urria S., Venegas A.;
RT "Serological response to Helicobacter pylori recombinant antigens in
RT Chilean infected patients with duodenal ulcer, non-ulcer dyspepsia and
RT gastric cancer.";
RL APMS 107:1069-1078 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1.
RX MEDLINE=22121498; PubMed=12125208;
RA Muller I., Medina-Selby A., Palacios J.L., Martinez P., Opazo P.,
RA Bruce E., Mancilla M., Valenzuela P., Yudelevich A., Venegas A.;
RT "Cloning and comparison of ten gene sequences of a Chilean H. pylori
RT strain with other H. pylori strains revealed higher variability for
RT Vaca and CagA virulence factors.";
RL Biol. Res. 35:67-84 (2002).
DR EMBL; AF479027; AAL86897.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SNR; Q8RNU5; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.

Query Match 51.8%; Score 611.5; DB 2; Length 238;
Best Local Similarity 52.7%; Pred. No. 2.5e-43;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEKGKILNYYEAVLISAHIMEEARGKKTAAELMQE 60

Qy 61 CMHFLKDEVMGPNVVDLGVETATPDGKLVTVNWPIDPEHFKAGEVKFGCDKDI 120
Db 61 GRTLKPDVMDGVASMIHEVGIEAMPDPDGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFPEANKALFKDREKAYGKRLDIPSGNLTIRIGAG 180
Db 120 INEGKAVSVKVNVDGPRVQIGSHFFFEVNRCLDFDREKTFGKRLDIASGTAVRFPFG 179

Qy 181 QTRKVQLIPLGSKKVGIMGNLVNNIADERHGHKALDKAKSHGF 224
Db 180 EEKSVELIDIGGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223

RESULT 5
Q93NJ1 HELPY
ID Q93NJ1_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373558; AAK69724.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SNR; Q93NJ1; 1-238.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR ProDom; PD002326; Urease_gamma_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26672 MW; EE82BF861D03BA50 CRC64;

Query Match 51.8%; Score 611.5; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 2.5e-43;
```


FT TURN 67 69
FT STRAND 70 70
FT TURN 72 73
FT HELIX 74 77
FT STRAND 80 87
FT TURN 88 89
FT STRAND 90 97
FT TURN 108 109
FT STRAND 111 111
FT STRAND 117 119
FT TURN 120 121
FT STRAND 128 133
FT STRAND 139 142
FT TURN 143 144
FT HELIX 147 149
FT TURN 152 153
FT STRAND 154 155
FT HELIX 158 161
FT TURN 162 163
FT STRAND 164 166
FT TURN 170 171
FT STRAND 173 176
FT TURN 178 179
FT STRAND 181 188
FT TURN 191 192
FT STRAND 194 195
FT TURN 198 199
FT STRAND 204 205
FT HELIX 208 221
FT TURN 222 222
FT TURN 224 225
SQ SEQUENCE 238 AA; 4E77328669CD9A2D CRC64;

Query Match 51.7%; Score 610.5; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 3e-43;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;
Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKKEGKIKLYVEAVALLISAHIMEEARAGKKSAAELMQE 60
Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTGLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKXAVSVKVNVDGDPVQIGSHFFHFFVNRVCLDFDREKTFGKRLDIASGTAVRPEPG 179
Qy 181 QTRKVLIPLGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223

RESULT 7
Q93NI3_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373570; AAK69736.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.

DR SMR; Q93NI3; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; P:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26586 MW; DFDBA17D016FC351 CRC64;
Query Match 51.7%; Score 610.5; DB 2; Length 238;
Best Local Similarity 53.1%; Pred. No. 3e-43;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKKEGKIKLYVEAVALLISAHIMEEARAGKKSAAELMQE 60
Qy 61 CMHFLKKDEVMGPNVDPDLGVEATFPDGTGLVTNNWPIEDPDEHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119
Qy 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFPEANKALKPDREKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKXAVSVKVNVDGDPVQIGSHFFHFFVNRVCLDFDREKTFGKRLDIASGTAVRPEPG 179
Qy 181 QTRKVLIPLGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDRQADNESKKTALHRAKERGF 223

RESULT 8
Q93NI5_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NI5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
GN Name=ureA;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chu C., Ou J.T.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF373568; AAK69734.1; -; Genomic_DNA.
DR HSSP; P14916; IE9Y.
DR SMR; Q93NI5; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; P:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR002019; Urease_beta.
DR InterPro; IPR008223; Urease_gamma.
DR InterPro; IPR002026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR PIRSF; PIRSF001225; Urease_gamma_beta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMs; TIGR00192; urease_beta; 1.
DR TIGRFAMs; TIGR00193; urease_gamma; 1.
SQ SEQUENCE 238 AA; 26539 MW; 251AD890042A4262 CRC64;

[illegible]

```

181 QTRKVLQIPIGGSKVVTGMGLVNNIADRHKKHAKDQKSHGF 224
: : : : : : : : : : : : : : : : : : : : : :
180 EEXSVELIDIGGNRRIRFGFVALVDQADNESKKIALHRAKERGF 223

RESULT 13
Q93NJ0_HELPY
ID Q93NJ0_HELPY PRELIMINARY; PRT; 238 AA.
AC Q93NJ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease protein Urea.
DE Name=ureA;
GN GN
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacter
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
```

Submitted (APR-2001) to the EMBL/GenBank/DBAJ databases.
DR EMBL; AF373560; AAK69726.1; -; Genomic_DNA.
DR HSSP; P14916; 1E9Y.
DR SMR; Q93NJ0; 1-238.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPRO002019; Urease_beta.
DR InterPro; IPRO08223; Urease_gammabeta.
DR InterPro; IPRO02026; Urease_gamma_reg.
DR Pfam; PF00699; Urease_beta; 1.
DR Pfam; PF00547; Urease_gamma; 1.
DR TIRSP; TIRSP001225; Urease_gammabeta; 1.
DR ProDom; PD002326; Urease_beta; 1.
DR ProDom; PD002319; Urease_gamma; 1.
DR TIGRFAMS; TIGR00192; urease_beta; 1.
DR TIGRFAMS; TIGR00193; urease_gamma; 1.
DR SEQUENCE 238 AA; 26629 MW; 5C082660AAC88469 CRC64;
SQ

```

181 QTRKVLQIPGGKKVIGMGLVNNIADERHKKHAKLSHGK 224
: ||| :||::: ||: ||| :|||
180 EEKSVELIDIGGRRIRFGFVALVDROQADNESKIALHRAKERGF 223

RESULT 15
Q84F76_HELPY
ID Q84F76_HELPY PRELIMINARY; PRT; 238 AA.
AC Q84F76;
DT 01-JUN-2003 (TREMBlurel. 24, Created)
DT 01-JUN-2003 (TREMBlurel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlurel. 26, Last annotation update)
DE Urease alpha (EC 3.5.1.5).
GN Name:urea;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]

```

```

AC  KARIM=iran-AF051;
RA  Karim Arzenani M., Mohammadi M.;
RL  Submitted (JAN-2003) to the EMBL/GenBank/DBDJ databases.
DR  EMBL; AY227442; AA034402.1; -; Genomic_DNA.
DR  HSSP; P14916; IE9Y.
DR  SNR; Q84F76; 1-238.
DR  GO; GO:0005737; C:Cytoplasm; IEA.
DR  GO; GO:0016787; F:Hydrolase activity; IEA.
DR  GO; GO:0016151; F:Nickel ion binding; IEA.
DR  GO; GO:0009039; F:Urease activity; IEA.
DR  GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR  InterPro; IPR002019; Urease_beta.
DR  InterPro; IPR008223; Urease_gamma_beta.
DR  InterPro; IPR002026; Urease_gamma_reg.
DR  Pfam; PF00699; Urease_beta; 1.
DR  Pfam; PF00547; Urease_gamma; 1.
DR  PRSF; PRSF001225; Urease_gamma_beta; 1.
DR  ProDom; PD002326; Urease_beta; 1.
DR  ProDom; PD002319; Urease_gamma; 1.
DR  TIGRFAMs; TIGR00192; urease_beta; 1.
DR  TIGRFAMs; TIGR00193; urease_gamma; 1.
KW  Hydrolase.
SQ  SEQUENCE      238 AA;  26491 MW;  6384770531FCB38C CRC64;
      Query Match      51.3%;  Score 605.5;  DB 2;  Length 238;
      Best Local Similarity 52.7%;  Pred. No. 8e-43;
      Matches 118;  Conservative 42;  Mismatches 63;  Indels 1;  Gaps 1;
      Qy  1  VKLTPQEKFLLYYAGEVARRKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
      Db  1  MKLTPKELDKMLHYAGELAKRKEGKIKLNVYEAVAILSAHIMEEARAGKKTAAELMQE 60
      Qy  61  CMHFLKDEVMVPGNVNVPDLGVGEATPDGTKLTVVNWPIEDPHFKAGEVKGCDKDIE 120
      Db  61  GRTLLKDDVMDGVASMIHEVGIEAMPDGTGLTVVHTPIEANGKLVPGEU-FLKNEDIT 119

```

```

QY      121 LNAGKEVIELEVINEGPKSLHVGSHFFPFEAKKALKDREKAIGKRLLUIPSGNIDRIKIGAG 180
Db      120 INEGKKAVSVKVKIVGDRPVQVQSGHLHFFEYNRCLDFEDREXTFGKRDLIASCTAVRFEPG 179

QY      181 QTRKVQLIPLGGSKKVI GMNGLVNNIADERHKHKALDKAKSHGF 224
Db      180 EKSVELIDIGGNRRITFGFNALVDROADNESKKTALHRAKERGF 223

Search completed: November 28, 2005, 08:19:44
Job time : 88.8136 secs
```

Search completed: November 28, 2005, 08:19:44
Job time : 88.8136 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:55:02 ; Search time 73.7204 Seconds
(without alignments)
1346.976 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADRRHKHALDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1181	100.0	226	5	Adj58238 Urease su
2	1178	99.7	226	5	Adj58250 UreaseXY
3	1174	99.4	226	5	Adj58241 UreaseXY
4	1174	99.4	226	5	Adj58247 UreaseXY
5	1170	99.1	226	5	Adj58244 UreaseXY
6	652	55.2	234	8	Adq37849 H. bizzo
7	613.5	51.9	238	4	AAU35875 Helicobac
8	610.5	51.7	238	2	AAR04579 Part of p
9	610.5	51.7	238	2	AAR07193 H. pylori
10	610.5	51.7	238	4	AAU35693 Helicobac
11	610.5	51.7	238	5	AAU49644 Urease A
12	610.5	51.7	245	4	AEBS5115 Hylcobac
13	607.5	51.4	238	8	ADS09178 H. pylori
14	603.5	51.1	238	8	ADM28643 Helicobac
15	602.5	51.0	238	2	AAR67375 H. pylori
16	596.5	50.5	238	5	AAU49645 Salmonell
17	595.5	50.4	238	2	AAR12515 A subunit
18	573.5	48.6	228	3	ABBS2559 Helicobac
19	573.5	48.6	228	5	ABU51893 Helicobac
20	572.5	48.5	237	2	AAR74336 Helicobac
21	572.5	48.5	237	2	AAW06729 H. felis
22	572.5	48.5	806	2	AAR67371 Urease A
23	562.5	47.6	222	5	AAU49647 Gerbil ur
24	559.5	47.4	219	5	AAU49646 Murine ur

25	559.5	47.4	224	5	AAU49648	Aam49648 Murine ur
26	502.5	42.5	811	6	ABU42650	Abu42650 Protein e
27	497.5	42.1	207	6	ABU42523	Abu42523 Protein e
28	493	41.7	779	6	ABU41210	Abu41210 Protein e
29	489.5	41.4	226	4	AAU33952	Aau33952 Staphyloc
30	477	40.4	207	6	ABU20285	Abu20285 Protein e
31	470.5	39.8	209	6	ABU40327	Abu40327 Protein e
32	470	39.8	261	6	ABU27583	Abu27583 Protein e
33	469.5	39.8	207	6	ABU17114	Abu17114 Protein e
34	455.5	38.6	228	6	ABU22153	Abu22153 Protein e
35	449.5	38.1	209	2	AAR67376	Aar67376 P. mirabi
36	443.5	37.6	1216	2	AAW14497	Aaw14497 Urease pr
37	439.5	37.2	840	8	ADJ66032	Adj66032 Coil form
38	439.5	37.2	840	9	AEA54210	Aea54210 C. ensifo
39	437	37.0	161	5	AEU50981	Aeu50981 Helicobac
40	429.5	36.4	270	2	AAR67377	Aar67377 Canavalia
41	368.5	31.2	837	8	ADG65522	Adg65522 C. posada
42	281.5	23.8	100	2	AAW37775	Aaw37775 Klebsiell
43	281.5	23.8	100	3	AAU81823	Aay81823 Klebsiell
44	281.5	23.8	100	4	AAU36152	Aau36152 Klebsiell
45	281.5	23.8	100	6	ABU32326	Abu32326 Protein e

ALIGNMENTS

RESULT 1
ADJ58238
ID ADJ58238 standard; protein; 226 AA.
XX
AC ADJ58238;
AC
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide X.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
FN EPI176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
N-PSDB; ADJ58237.
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
Claim 8; SEQ ID NO 2; 76pp; English.
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an urease
X subunit polypeptide of the invention.

Query Match 100.0%; Score 1181; DB 5; Length 226;
SQ Sequence 226 AA;

```
Best Local Similarity 100.0%; Pred. No. 6.8e-116;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAEGKRLDIPSGNTLRIGAG 120
DB 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAEGKRLDIPSGNTLRIGAG 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVIWMGLVNNIADBRHKHKKALDKAKSHGFIK 226
DB 181 QTRKVQLIPLGGSKKVIWMGLVNNIADBRHKHKKALDKAKSHGFIK 226

RESULT 2
ADJ58250
ID ADJ58250 standard; protein; 226 AA.
XX
AC ADJ58250;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #7.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58249.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 14; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;
XX
Query Match 99.7%; Score 1178; DB 5; Length 226;
Best Local Similarity 99.6%; Pred. No. 1.4e-115;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
QY 181 QTRKVQLIPLGGSKKVIWMGLVNNIADBRHKHKKALDKAKSHGFIK 226
DB 181 QTRKVQLIPLGGSKKVIWMGLVNNIADBRHKHKKALDKAKSHGFIK 226

RESULT 3
ADJ58241
ID ADJ58241 standard; protein; 226 AA.
XX
AC ADJ58241;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseXY subunit #1.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58240.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX
PS Disclosure; SEQ ID NO 5; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;
XX
Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 3.7e-115;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 VKLTPKEQKFLYYAGEVARRKKAEGKLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAEGKRLDIPSGNTLRIGAG 120
DB 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTGTLVTNNWPIEPDEHFKAEGKRLDIPSGNTLRIGAG 120
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
DB 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDEKAYGKRLDIPSGNTLRIGAG 180
```

```
Db 121 LNAGKEVTELVNEGPKSLHVGSHFFETNKAALFDRKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226

RESULT 4
ADJ58247
ID ADJ58247 standard; protein; 226 AA.
XX
AC ADJ58247;
XX
DT 06-MAY-2004 (first entry)
DE UreaseXY subunit #5.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58246.
XX
PS Disclosure; SEQ ID NO 11; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;

Query Match 99.4%; Score 1174; DB 5; Length 226;
Best Local Similarity 99.1%; Pred. No. 3.7e-115;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAAGEVFGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAAGEVFGCDKDIE 120
Qy 121 LNAGKEVTELVNEGPKSLHVGSHFFETNKAALFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELVNEGPKSLHVGSHFFETNKAALFDRKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
```

```
RESULT 5
ADJ58244
ID ADJ58244 standard; protein; 226 AA.
XX
AC ADJ58244;
XX
DT 06-MAY-2004 (first entry)
DE UreaseXY subunit #3.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
DR WPI; 2002-124384/17.
DR N-PSDB; ADJ58243.
XX
PS Disclosure; SEQ ID NO 8; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
SQ Sequence 226 AA;

Query Match 99.1%; Score 1170; DB 5; Length 226;
Best Local Similarity 98.7%; Pred. No. 9.8e-115;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAELMEE 60
Qy 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAAGEVFGCDKDIE 120
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAAGEVFGCDKDIE 120
Qy 121 LNAGKEVTELVNEGPKSLHVGSHFFETNKAALFDRKAYGKRLDIPSGNTLRIGAG 180
Db 121 LNAGKEVTELVNEGPKSLHVGSHFFETNKAALFDRKAYGKRLDIPSGNTLRIGAG 180
Qy 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226
Db 181 QTRKQVLIPLGSKKVIQGNGLVNNIADERHKHAKLDKAKSHGFIK 226

RESULT 6
ADQ37849
ID ADQ37849 standard; protein; 234 AA.
XX
```

AC ADQ37849;
XX 07-OCT-2004 (first entry)
XX H. bizzozeronii urea polypeptide.
XX Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureA; Helicobacter bizzozeronii infection;
KW antibacterial; enzyme.
XX Helicobacter bizzozeronii.
OS US2004142343-A1.
XX 22-JUL-2004.
XX 12-AUG-2003; 2003US-00639273.
XX 16-AUG-2002; 2002US-0404337P.
XX (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
BA (ZHUJ/) ZHU J.
PI Chang Y, Simpson KW, Zhu J;
XX WPI; 2004-533502/51.
DR N-PSDB; ADQ37848.
XX Novel isolated nucleic acid molecule having urease gene cluster, and
PT conferring on Helicobacter bizzozeronii ability to produce urease, useful
PT as vaccine for preventing disease in mammals infected by H.bizzozeronii.
PS Claim 23; SEQ ID NO 3; 40pp; English.
XX The invention relates to an isolated nucleic acid molecule conferring on
CC Helicobacter bizzozeronii an ability to produce urease, where the nucleic
CC acid molecule is a urease gene cluster comprising at least one urease
CC structural gene and at least one urease accessory gene. The nucleic acid
CC molecule is chosen from ureA, ureB, ureF, ureG, ureH and ureI. The
CC invention also relates to an isolated protein encoded by the nucleic
CC acid, a vaccine for preventing onset of disease in mammals infected by H.
CC bizzozeronii comprising a nucleic acid and a carrier, and an isolated
CC antibody or its binding portion raised against the nucleic acid. The
CC nucleic acids, proteins and antibodies are useful for vaccinating mammals
CC against onset of disease caused by infection of H. bizzozeronii, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizzozeronii in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizzozeronii is
CC present in the sample. This sequence represents the H. bizzozeronii urea
XX polypeptide.
XX Sequence 234 AA;
Query Match 55.2%; Score 652; DB 8; Length 234;
Best Local Similarity 57.3%; Pred. No. 5e-60;
Matches 129; Conservative 39; Mismatches 55; Indels 2; Gaps 2;
QY 1 VKLTPKEQKFLYYAGEVARKKASGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGELAKKRGKNGVKLNLYTEAVALISAHVMEERAGKKSVDLMQE 60
QY 61 CMHFLKKDEVMPCGVGNVPDLGVEATFPDGTCKLVTNNWPIEDPHFKA-CEVRFQGDCKI 119
DB 61 GRTLLKADDVMPGVARHIEHVGIANFPDGTCKLVTHTPVEDGGHKLAPCEVILK-NEDI 119
QY 120 ELNAGKEVTEVTNPGKSLHSHFHPPEANKALKPDRKAYGKRLDTPSGNTRLIGA 179
DB 120 TLNAGKQATTLVHNKGDPRVPQVQSGSHFHPPEVNVKLLFEDREKAYGKRLDIASGTAVRFP 179

QY 180 GQTRKVLQIPLGGSKKVGWGLVNNIADSRHKHKALDKAKSHGF 224
DB 180 GEKKTVELIQIGGNRIYGFNSLVDRQADTDGKKALKRAKEHGF 224
RESULT 7
AAU35875
ID AAU35875 standard; protein; 238 AA.
XX AAU35875;
XX 14-FEB-2002 (first entry)
XX Helicobacter pylori cellular proliferation protein #188.
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
XX Helicobacter pylori.
XX WO200170955-A2.
PN 27-SEP-2001.
XX 21-MAR-2001; 2001WO-US0009180.
XX 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 23-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX (ELIT-) ELITRA PHARM INC.
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;
PI Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
DR N-PSDB; AAS53734.
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
PS Example 3; SEQ ID NO 11468; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 238 AA;
Query Match 51.9%; Score 613.5; DB 4; Length 238;
Best Local Similarity 53.1%; Pred. No. 6e-56;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;
QY 1 VKLTPKEQKFLYYAGEVARKKASGLKLNQPEAIAISAHIMDEARRGKKTVAQLMEE 60

Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYEAVALLSAHIMEEARAGKKTAAELMQE 60
QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTCLVTNWPIDEPDEHFKAQGVKFGCDKDIE 120
Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTCLVTHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFKDREKAYGKRDLIPSGNTLRIGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRCLDFDREKTFGKRDLIASGTAVRFBPG 179
QY 181 QTRKVLIPLGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EEKSVELIDIGNRRIFGFNALVDROADNESKKIALHRAKERGF 223

RESULT 8

AA04579

ID AAR04579 standard; protein; 238 AA.

XX AAR04579;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 21-SEP-1990 (first entry)

XX Part of protein with urease activity.

XX Urease; probe.

XX Helicobacter pylori.

OS WO9004030-A.

XX 19-APR-1990.

XX 06-OCT-1988; 88FR-00013135.

XX 06-OCT-1988; 88FR-00013135.

XX (INSP) INST PASTEUR.

PA (INRM) INSERM INST NAT SANTS & RECH MED.

XX Labigne A;

XX WPI; 1990-147844/19.

DR N-PSDB; AARQ4328.

XX New nucleotide sequences encoding Campylobacter pylori-ureaseants - and
PT derived vectors, transformants, protein, antibodies and probes, useful in
PT diagnosis, treatment and prevention of infections.

XX Claim 11; Page 34; 47pp; French.

PS The protein can be used for the prodn. of antibodies and to prepare

CC vaccines for the prevention /treatment of C. pylori infections. See also

CC AAR04578-81. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27

CC -AUG-2003 to correct OS field.)

XX Sequence 238 AA;

Query Match

Best Local Similarity 51.7%; Score 610.5; DB 2; Length 238;

Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60

Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYEAVALLSAHIMEEARAGKKTAAELMQE 60

QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTCLVTNWPIDEPDEHFKAQGVKFGCDKDIE 120

Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTCLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFKDREKAYGKRDLIPSGNTLRIGAG 180
Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRCLDFDREKTFGKRDLIASGTAVRFBPG 179
QY 181 QTRKVLIPLGSKKVGIMGLVNNIADERHKHAKLDKAKSHGF 224
Db 180 EEKSVELIDIGNRRIFGFNALVDROADNESKKIALHRAKERGF 223

RESULT 9

AA07193

ID AAW07193 standard; protein; 238 AA.

XX AAW07193;

DT 16-OCT-2003 (revised)

DT 11-FEB-1997 (first entry)

XX H. pylori urease A subunit.

XX Urease; ureA gene; ureB gene; vaccine.

XX Helicobacter pylori; strain CPM630.

XX WO9633732-A1.

XX 31-OCT-1996.

XX 25-APR-1996; 96WO-US005800.

XX 28-APR-1995; 95US-00431041.

PR 06-DEC-1995; 95US-00568122.

XX (ORAV-) ORAVAX INC.

XX Lee CK, Monath TP, Ackerman SK, Thomas WD, Sonan G, Kleanthous H;
PI Weltzin RA, Pappo J, Ermak T, Guirakhoo F, Bhagat H, Sussman I;

XX WPI; 1996-497373/49.

XX N-PSDB; AAT44351.

XX Vaccine for inducing mucosal response to Helicobacter - contg. multimeric
PT urease complex and pref. an antibiotic, anti-secretory agent or bismuth
PT salt.

XX Disclosure; Page 70-71; 98pp; English.

XX Urease A (AAW07193) and B (AAW07194) subunits of Helicobacter pylori are
CC encoded by the ureA + ureB gene locus of clinical isolate CPM630.
CC Vectors, e.g. pORV214 (see also AAT44351), carrying the gene locus can be
CC used to transform host cells for the large-scale prodn. of recombinant,
CC enzymatically inactive, multimeric urease. The urease complex induces a
CC mucosal immune response that can treat or prevent Helicobacter, esp. H.
CC pylori, gastroduodenal infection. (Updated on 16-OCT-2003 to standardise
CC OS field)

XX Sequence 238 AA;

Query Match

Best Local Similarity 51.7%; Score 610.5; DB 2; Length 238;

Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIYISAHIMDEARRGKKTVAOLMEE 60

Db 1 MKLTPKELDKMLHYAGELARRKKEGKILNYYEAVALLSAHIMEEARAGKKTAAELMQE 60

QY 61 CMHFLKDEVMFGVGNMVPDLGVEATFPDGTCLVTNWPIDEPDEHFKAQGVKFGCDKDIE 120

Db 61 GRTLLKPDVDMGVASMIHEVGIEAMFPDGTCLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFKDREKAYGKRDLIPSGNTLRIGAG 180

Db 120 INEGKAVSVKXNVGDRPVQIGSHFHFPEVNRCLDFDREKTFGKRDLIASGTAVRFBPG 179

CC included are a gene which encodes the urease fragment, a vaccine against
 CC H. pylori comprising the urease fragment, a transformed cell transformed
 CC by introducing the gene, an antibody which recognises the urease fragment
 CC and transforming a plant by introducing the gene into the plant and
 CC making it express. The urease fragment is useful for producing blood
 CC serum containing antibodies which suppresses proliferation of H. pylori
 CC and which suppresses urease activity of H. pylori which involves
 CC immunising a living organism (e.g. a human, cow or hen) with the urease
 CC fragment. The fragment may be purified from cow's milk or hen eggs from
 CC transgenic animals expressing the fragment. H. pylori is thought to be a
 CC cause of gastric and peptic ulcers, gastritis and may be implicated in
 CC stomach cancer. The urease enzyme is thought to be responsible for
 CC maintaining the bacterium in the acid conditions of the stomach. The
 CC present sequence is the H. pylori urease alpha subunit.

XX Sequence 238 AA;

Query Match 51.4%; Score 607.5; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 2.6e-55;
 Matches 117; Conservative 44; Mismatches 62; Indels 1; Gaps 1;
 QY 1 VKLTPEQEKFLLYAGVAVRKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
 DB 1 MKLTPKELDKMLHYAGELAKKREKIKNTVEAVALISAHIMEEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMGVMGNNVDPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
 DB 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 DB 120 INEGKAVSVKVNVDGPRVQIGSHFFHFFPEVNRCLDFDREKTFGKRLDIAAGTAVRFPFG 179
 QY 181 QTRKVLQPLGSKKVIQNGNLVNNIADERHKHKKALDKAKSHGF 224
 DB 180 EKSVELIDIGNRRIFGPNALVDQADNESKKIALHRAKERGF 223

RESULT 14

ADM28643
 ID ADM28643 standard; protein; 238 AA.

AC ADM28643;

XX 20-MAY-2004 (first entry)

DT Helicobacter pylori urease alpha subunit protein SeqID 2.

DE immunogenic; urease; vaccine; passive immunisation; diagnostic;

KW antibacterial; antibody.

XX Helicobacter pylori.

OS JP2004041084-A.

PN 12-FEB-2004.

XX 11-JUL-2002; 2002JP-00203221.

XX 11-JUL-2002; 2002JP-00203221.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2004-162021/16.

DR New peptide useful as a vaccine against Helicobacter pylori induces an

PT antibody response against the bacterial urease.

XX Disclosure; SEQ ID NO 2; 19pp; Japanese.

XX This invention relates to a novel immunogenic peptide that induces

CC production of antibodies against the Helicobacter pylori (H. pylori)

CC urease protein. Specifically, it refers to a peptide antigen capable of

CC inducing an antibody response, such that it can be used to develop a
 CC vaccine against H. pylori. The present invention describes generating
 CC monoclonal antibodies against the urease using hybridoma techniques that
 CC are useful for passive immunisation and diagnostic techniques.
 CC Furthermore, the antibacterial peptide can be used to immunise an animal
 CC in order to produce antibodies that are present within a functional food
 CC such as milk or eggs. This polypeptide sequence is the H. pylori urease
 CC alpha subunit protein of the invention.

XX Sequence 238 AA;

Query Match 51.1%; Score 603.5; DB 8; Length 238;
 Best Local Similarity 52.2%; Pred. No. 6.8e-55;
 Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLLYAGVAVRKAEGLKLNQPEAIYISAHIMDEARRGKKTVAQLMEE 60
 DB 1 MKLTPKELDKMLHYAGELAKKREKIKNTVEAVALISAHIMEEARAGKKTAAELMQE 60
 QY 61 CMHFLKKDEVMGVMGNNVDPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
 DB 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTCLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
 QY 121 LNAGKEVTELVNTEGPKSLHVGSHFFHFFPEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
 DB 120 INEGKAVSVKVNVDGPRVQIGSHFFHFFPEVNRCLDFDREKTFGKRLDIAAGTAVRFPFG 179
 QY 181 QTRKVLQPLGSKKVIQNGNLVNNIADERHKHKKALDKAKSHGF 224
 DB 180 EKSVELIDIGNRRIFGPNALVDQADNESKKIALHRAKERGF 223

RESULT 15

AAR67375
 ID AAR67375 standard; protein; 238 AA.

AC AAR67375;

XX 25-MAR-2003 (revised)

DT 22-JUN-1995 (first entry)

XX H. pylori ureA urease.

XX Urease; ureA gene; immunogen; vaccine; diagnostic; Helicobacter felis.

OS Proteus mirabilis.

XX WO9426901-A1.

PD 24-NOV-1994.

XX 19-MAY-1994; 94WO-EP001625.

XX 19-MAY-1993; 93EP-00401309.

XX 19-NOV-1993; 93WO-EP003259.

XX (INSP) INST PASTEUR.

XX (INRM) INST NAT SANTE & RECH MEDICALE.

XX Labigne A, Suerbaum S, Ferrero R, Thiberge J;

XX WPI; 1995-006797/01.

PT DNA from Helicobacter pylori and Helicobacter felis - used to develop

PT prods. for detection, treatment and prevention of Helicobacter infection.

XX Disclosure; Fig 4i-iii; 168pp; English.

XX The sequence of the Helicobacter felis urease ureA gene product (given in

CC AAR67371) was compared to ureases of Helicobacter pylori (AAR67375),

CC Proteus mirabilis (AAR67376) and jack bean urease (AAR67377) and regions

CC of homology were identified. (Updated on 25-MAR-2003 to correct PN

```
XX
SQ      Sequence 238 AA;
      Query Match      51.0%; Score 602.5; DB 2; Length 238;
      Best Local Similarity 52.2%; Pred.No. 8.7e-55;
      Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

QY      1 VKLTPKEQKFLLYAGEVARKKAEGKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
      :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 MKLTPKELDKLMLHYAGELAKRKEGKILNYVEAVALISAHIMEEARAGKKTAAELMQE 60

QY      61 CMHFLKKDEVMPGVGNMVPDLGVEATPPDGTKLVTWNWPIEPDEHFKAGEVKFGCDKDIE 120
      || : || : || : || : || : || : || : || : || : || : || : || : ||
Db      61 GRTLLKPDVVMDGVASMIHEVGIEAMPDGTKLVTHTPIEANGKLVPGEL-FLKNEDIT 119

QY      121 LNAGKEVTELEVTNEGPKSLHVGSHPHFFPEANKALKFDREKAYCKRLDIPSGNTRLRICAG 180
      : || : || : || : || : || : || : || : || : || : || : || : ||
Db      120 INEGKAVSVKXNVGDRPVQIGSHPHFFVFNRCCLDPDREKTFHRLUDIASGTAVRFPFG 179

QY      181 QTRKVLIPLGSGKKVIGMNLVNNIADERRHKIKALDKAKSHGF 224
      : : || : || : : : : || : || : || : || : || : || : ||
Db      180 EEKSVELLIDIGGNRRIFGFGNALVDRQADNESKKIALHRAKERGF 223
```

Search completed: November 28, 2005, 08:14:34
Job time : 76.7204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:11:38 ; Search time 63.7582 Seconds
(without alignments)
1481.056 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHKKALDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1181	100.0	226	3	US-09-904-994B-2
2	1178	99.7	226	3	US-09-904-994B-14
3	1174	99.4	226	3	US-09-904-994B-5
4	1170	99.4	226	3	US-09-904-994B-11
5	1170	99.1	226	3	US-09-904-994B-8
6	652	55.2	234	4	US-10-639-273-3
7	622	52.7	234	4	US-10-639-273-34
8	613.5	51.9	238	3	US-09-815-242-11468
9	613.5	51.9	238	4	US-10-335-977-8629
10	610.5	51.7	238	3	US-09-815-242-11286
11	594	50.3	237	4	US-10-639-273-36
12	576.5	48.8	228	4	US-10-335-977-8628
13	573.5	48.6	228	4	US-10-012-819-154
14	566.5	48.0	237	4	US-10-639-273-35
15	540.5	45.8	225	4	US-10-639-273-37
16	502.5	42.5	811	4	US-10-282-122A-70574
17	497.5	42.1	207	4	US-10-282-122A-70447
18	493	41.7	779	4	US-10-282-122A-69134
19	489.5	41.4	207	3	US-09-815-242-5448
20	477	40.4	226	4	US-10-282-122A-48209
21	470.5	39.8	209	4	US-10-282-122A-68251
22	470	39.8	261	4	US-10-282-122A-55507
23	469.5	39.8	207	4	US-10-282-122A-45038
24	455.5	38.6	228	4	US-10-282-122A-50077
25	439.5	37.2	840	4	US-10-621-833-7
26	439.5	37.2	840	5	US-10-731-877-1
27	439.5	37.2	840	6	US-11-046-271-1

28	426	36.1	843	4	US-10-437-963-124057	Sequence 124057,
29	424.5	35.9	171	4	US-10-335-977-8627	Sequence 8627, Ap
30	384.5	32.6	227	4	US-10-156-761-10253	Sequence 10253, A
31	381.5	32.3	837	4	US-10-424-599-254635	Sequence 254635,
32	368.5	31.2	837	4	US-10-418-962-2	Sequence 2, Appl1
33	291	24.6	227	4	US-10-425-115-201340	Sequence 201340,
34	281.5	23.8	100	3	US-09-815-242-11745	Sequence 11745, A
35	281.5	23.8	100	4	US-10-156-761-14633	Sequence 14633, A
36	281.5	23.7	100	4	US-10-282-122A-60250	Sequence 60250, A
37	279.5	23.7	121	4	US-10-282-122A-49772	Sequence 49772, A
38	278.5	23.6	100	4	US-10-282-122A-62539	Sequence 62539, A
39	278.5	23.5	100	4	US-10-282-122A-64616	Sequence 64616, A
40	277.5	23.5	100	4	US-10-282-122A-56822	Sequence 56822, A
41	277.5	23.5	100	4	US-10-282-122A-56846	Sequence 56846, A
42	267.5	22.7	100	3	US-09-815-242-11047	Sequence 11047, A
43	267.5	22.7	100	4	US-10-282-122A-58211	Sequence 58211, A
44	267.5	22.7	100	4	US-10-282-122A-69398	Sequence 69398, A
45	265.5	22.5	100	4	US-10-282-122A-44442	Sequence 44442, A

ALIGNMENTS

RESULT 1

US-09-904-994B-2
; Sequence 2, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-2

Query Match 100.0%; Score 1181; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.4e-114;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE	60
Db	1	VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE	60
Qy	61	CMHFLKKDEVMPGVGNMVPDLGVETATPDGKLVTVNWPIDPDEHFHFKAGEVFGCDKQIE	120
Db	61	CMHFLKKDEVMPGVGNMVPDLGVETATPDGKLVTVNWPIDPDEHFHFKAGEVFGCDKQIE	120
Qy	121	LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKPDREKAYKRLDIPSGNTLRIGAG	180
Db	121	LNAGKEVTELEVTNEGPKSLHVGSHFFHFEANKALKPDREKAYKRLDIPSGNTLRIGAG	180
Qy	181	QTRKVLIPLGSKKVGIMGNLVNNIADERHKHKKALDKAKSHGFIK	226
Db	181	QTRKVLIPLGSKKVGIMGNLVNNIADERHKHKKALDKAKSHGFIK	226

RESULT 2

US-09-904-994B-14
; Sequence 14, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-14

Query Match      99.7%; Score 1178; DB 3; Length 226;
Best Local Similarity 99.6%; Pred. No. 7.1e-114;
Matches 225; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
Db 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||

RESULT 3
US-09-904-994B-5
; Sequence 5, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-5

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 1.8e-113;
Matches 224; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
Db 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||

RESULT 4
US-09-904-994B-11
; Sequence 11, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-11

Query Match      99.4%; Score 1174; DB 3; Length 226;
Best Local Similarity 99.1%; Pred. No. 1.8e-113;
Matches 224; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
Db 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||

RESULT 5
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 4.8e-113;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
Db 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||

RESULT 6
US-09-904-994B-8
; Sequence 8, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-8

Query Match      99.1%; Score 1170; DB 3; Length 226;
Best Local Similarity 98.7%; Pred. No. 4.8e-113;
Matches 223; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
   |||||
Db 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAELMEE 60
   |||||
QY 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
Db 61 CMHFLKKDEVMPGVGNMVPDLGVEATFPDGTCLVTNNWPIEPDEHFKAGEVKGCDKDIE 120
   |||||
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
Db 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFFEANKALKFDREKAYGKRLDIPSGNTLRIGAG 180
   |||||
QY 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
Db 181 QTRKVQLIPLGGSKKVIQGNGLVNNIADERRHKHAKLDKAKSHGF 226
   |||||
```

```
US-10-639-273-3
; Sequence 3, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; PRIOR FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-3

Query Match 55.2%; Score 652; DB 4; Length 234;
Best Local Similarity 57.3%; Pred. No. 3.9e-59;
Matches 129; Conservative 39; Mismatches 55; Indels 2; Gaps 2;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMER 60
DB 1 MKLTPKELDKMLHYAGELARKKANGVKLNTEAVALLSAHMEEARAGKKSVAOLMQE 60

QY 61 CMHFLKKDEVMVPGVGNMVPDLGVEATFPDGTGKLVTVNWPIEP-GEVKFGCDKI 119
DB 61 GRTLLKADDVMVGAHMIHEVGIEAGFPDGTGKLVTHITPVEAGSKLAPGEVILK-NEDI 119

QY 120 ELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGA 179
DB 120 ELNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGA 179

QY 180 GQTRKVQLIPGGSKKVGIMGNLVNNIADDERHKHKALDKAKSHGF 224
DB 180 GEKTVELIDIGGNKRIYGFNALVDROADHDGKKLAKRAKEKHF 224

RESULT 7
US-10-639-273-34
; Sequence 34, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaqian
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; PRIOR FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-10-639-273-34

Query Match 52.7%; Score 622; DB 4; Length 234;
Best Local Similarity 54.7%; Pred. No. 5.2e-56;
Matches 123; Conservative 41; Mismatches 59; Indels 2; Gaps 2;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGKKTVAOLMER 60
DB 1 MKLTPKELDKMLHYAGELAKORKAKGIKLNYTEAVALLSAHMEEARAGKKSVAOLMQE 60

US-09-815-242-11468
; Sequence 11468, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11468
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11468

Query Match 51.9%; Score 613.5; DB 3; Length 238;
Best Local Similarity 53.1%; Pred. No. 4.1e-55;
Matches 119; Conservative 42; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLYYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAOLMEE 60
DB 1 MKLTPKELDKMLHYAGELARKKEGKILKNYVEAVALISAHIMEEARAGKKTAAELMQE 60

QY 61 CMHFLKKDEVMVPGVGNMVPDLGVEATFPDGTGKLVTVNWPIEPDPEHFKAGEVKFGCDKIE 120
DB 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119

QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKRLDIPSGNTLRIGAG 180
DB 120 INEGKKAASVKNVGNVDRPVOIGSHFHFVFNRCCLDFDREKTFGKRLDIASGTAVRFG 179

QY 181 QTRKVQLIPGGSKKVGIMGNLVNNIADDERHKHKALDKAKSHGF 224
DB 180 EEEKVELIDIGGNRIYGFNALVDROADNENSKIALHRAKERGF 223
```


; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 237
; TYPE: PR1
; ORGANISM: Helicobacter pylori
US-10-639-273-36

Query Match 50.3%; Score 594; DB 4; Length 237;
Best Local Similarity 52.2%; Pred. No. 4.4e-53;
Matches 117; Conservative 43; Mismatches 62; Indels 2; Gaps 2;

Qy 1 VKLTPEQKFLYYAGEVARKEKASGLKLNQPEATAYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQE 60

Qy 61 CMHFLKKDEVMPGVGNVVDLGEATFPDGTGLVTNWPIDEPDEHFHFKAGEVFGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 119

Qy 121 LNAGKEVTELVNNEGPKSLHVGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAG 180
Db 120 INEGKAVSVKVPVGDPRVQIGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAG 180

Qy 181 QTRKVOLIPLGSKVKVGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 180 EKSVELIDIGNRRIFGFNALVDROADNESK-IALHRAKERGF 222

RESULT 12
US-10-335-977-8628
; Sequence 8628, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandregouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 8628:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...228
; SEQUENCE DESCRIPTION: SEQ ID NO: 8628:
US-10-335-977-8628

Query Match 48.8%; Score 576.5; DB 4; Length 228;
Best Local Similarity 52.6%; Pred. No. 2.7e-51;
Matches 112; Conservative 40; Mismatches 60; Indels 1; Gaps 1;

Qy 12 LLYYAGEVARKRKAEGKLNQPEAIYISAHIMDEARRGKKTVAQLMEECMHFLKKDEV 71
Db 2 MLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQEGRTLLKPDV 61

Qy 72 PGVGNVVDLGEATFPDGTGLVTNWPIDEPDEHFHFKAGEVFGCDKDIELNAGKEVTELE 131
Db 62 DGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDITINEGKAVSVK 120

Qy 132 VTNEGPKSLHVGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAGOTRKKVOLIPG 191
Db 121 VKWGDPRVQIGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAGOTRKKVOLIPG 191

Qy 192 GSKKVIKMGVGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 181 GNRRIFGFNLVDROADNESKIALHRAKERGF 213

RESULT 13
US-10-012-819-154
; Sequence 154, Application US/10012819
; Publication No. US20030017478A1
; GENERAL INFORMATION:
; APPLICANT: Legrain, Pierre
; APPLICANT: Selig, Luc
; APPLICANT: Rain, Jean-Christophe
; TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
; FILE REFERENCE: B5053
; CURRENT APPLICATION NUMBER: US/10/012,819
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EP 99401066.8
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 228
; TYPE: PR1
; ORGANISM: Helicobacter pylori
US-10-012-819-154

Query Match 48.6%; Score 573.5; DB 4; Length 228;
Best Local Similarity 52.1%; Pred. No. 5.6e-51;
Matches 111; Conservative 41; Mismatches 60; Indels 1; Gaps 1;

Qy 12 LLYYAGEVARKRKAEGKLNQPEAIYISAHIMDEARRGKKTVAQLMEECMHFLKKDEV 71
Db 2 MLHYAGELAKRKEGKIKLNYVEAVRLISAHIMEEARRGKKTAAELMQEGRTLLKPDV 61

Qy 72 PGVGNVVDLGEATFPDGTGLVTNWPIDEPDEHFHFKAGEVFGCDKDIELNAGKEVTELE 131
Db 62 DGVASMIHEVGIEAMFPDGTGLVTHTPIEANGKLVPGEL-FLKNEDITINEGKAVSVK 120

Qy 132 VTNEGPKSLHVGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAGOTRKKVOLIPG 191
Db 121 VKWGDPRVQIGSHFFHFEVNRCLDPDRKAYGKRLDIPSGNTLRIGAGOTRKKVOLIPG 191

Qy 192 GSKKVIKMGVGNGLVNNIADERHKHKKALDKAKSHGF 224
Db 181 GNRRIFGFNLVDROADNESKIALHRAKERGF 213

Result No.	Score	Query Match	Length	DB	ID	Description	
1	247	20.9	111	1	US-10-793-626-116	Sequence 116, App	
2	212	18.0	106	1	US-10-485-517-181	Sequence 181, App	
3	84.5	7.2	247	1	US-10-793-626-2006	Sequence 2006, Ap	
4	77.5	6.6	259	1	US-10-510-386-108	Sequence 108, App	
5	76.5	6.5	428	1	US-10-689-742-50	Sequence 50, Appl	
6	73.5	6.2	429	1	US-10-821-234-1089	Sequence 1089, Ap	
7	72	6.1	418	1	US-10-858-730-6	Sequence 6, Appli	
8	71	6.0	311	1	US-10-793-626-1248	Sequence 1248, Ap	
9	70	5.9	329	1	US-10-793-626-1460	Sequence 1460, Ap	
10	70	5.9	437	1	US-10-821-234-1608	Sequence 1608, Ap	
11	69	5.8	501	1	US-10-793-626-244	Sequence 244, App	
12	68.5	5.8	1992	7	US-11-013-759-3	Sequence 3, Appli	
13	68.5	5.8	1992	7	US-11-013-759-13	Sequence 13, Appl	
14	68.5	5.8	2047	7	US-11-013-759-4	Sequence 4, Appli	
15	68.5	5.8	2047	7	US-11-013-759-7	Sequence 7, Appli	
16	67	5.7	932	7	US-11-017-550-65	Sequence 65, Appl	
17	67	5.7	1279	1	US-10-793-626-3188	Sequence 3188, Ap	
18	66.5	5.6	567	1	US-10-793-626-3184	Sequence 3184, Ap	
19	66.5	5.6	989	1	US-10-793-626-2394	Sequence 2394, Ap	
20	66	5.6	600	1	US-10-131-826A-462	Sequence 462, App	
21	65.5	5.5	328	1	US-10-821-234-1462	Sequence 1462, Ap	
22	65.5	5.5	1015	1	US-10-957-569-51	Sequence 51, Appl	
23	65.5	5.5	1168	1	US-10-509-422-2	Sequence 2, Appli	
24	65	5.4	397	1	US-10-485-517-172	Sequence 172, App	
25	64	5.4	302	1	US-10-793-626-2170	Sequence 2170, Ap	

```
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 181
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-181

Query Match      18.0%; Score 212; DB 1; Length 106;
Best Local Similarity 46.9%; Pred. No. 2.4e-14;
Matches 38; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

QY 138 KSLHVGSHHFFPEANKALKFDEKAYGKRLDIPSGNTLRIGAGQTRKVLIPLGSKKVI 197
DB 2 RPIQVGSFHFYFANAALDFEREMAYGKHLDPAGAAVRFEPGDKKEVOLVVEYAGKRTIF 61
QY 198 GNGLVNNIADERHKHAKLDK 218
DB 62 GPRGMWNGPIDESRVIRPTDE 82

RESULT 3
US-10-793-626-2006
; Sequence 2006, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2006
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2006

Query Match      7.2%; Score 84.5; DB 1; Length 247;
Best Local Similarity 23.8%; Pred. No. 0.14;
Matches 48; Conservative 34; Mismatches 69; Indels 51; Gaps 12;

QY 44 MDEARRGK-----KTVAQLMEECHFLKDKDEVMPGVGNMVPDLGVVEATFPDGTGL-- 93
DB 73 LEQARQGLAIWDHMLHTIEQPREELSAYPKVVTM----SINPKIRDVIGPGGKKINE 128
QY 94 ----VTVNWPIPDHFHFKAGEVKGKCDKDIELNAGKEVTELEVNTGPKSLHVGSHHFF 149
DB 129 IIDETGVKLDIEQD-----GTIFIGAVDAQMINRAKEIIE-DITREA---EVGVVH-- 176
QY 150 EANKALKFDEKAYGKRLDI-PSGNTL-----RIGAGQTRKVO-LIPLGSKKV-----IGM 199
DB 177 -----AKVKRIEYKGFVSEFLFGKQALHLHSIQISQERINKVEDVLKIGDTIEVKITEIDK 231
QY 200 NGLVNNIADERHKHAKLDAKS 221
DB 232 QGRVN-----ASHKVLEQSKN 247

RESULT 4
US-10-510-386-108
; Sequence 108, Application US/10510386
```

```
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Oisen, Peter Bjørke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 108
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-108

Query Match      6.6%; Score 77.5; DB 1; Length 259;
Best Local Similarity 21.7%; Pred. No. 0.71;
Matches 38; Conservative 36; Mismatches 54; Indels 47; Gaps 9;

QY 46 EARRGKKTVAQLMEECHFLKDKDEVMPGVGNMVPDLGVVEATFPDGTKLVTNNPIEDPH 105
DB 95 QKEKDKQSLKKIQEQVNRFFIKKNLQKQVNTKLTDEGLLS-----IEDNIF 141
QY 106 FXAGEVKGKCDKDIELNAGKEVTELEVNTGPKSLHVGSHF-----HFFEANKAL--- 155
DB 142 FDSGKAIR-QQDIPL--AKEVSDLLVNP-PRNIVISGHTDNPVIRNSQPKSNWHL SVM 197
QY 156 -----KFDRE-----KAYGRKLDIPSGNTLRIGAGQTRKVO--LIPLG 191
DB 198 RAVNFMGLLIENPKLDKIPFSAKGYGEFKPIASNDT-EEGRKNRRVEILLIPIG 251

RESULT 5
US-10-689-742-50
; Sequence 50, Application US/10689742
; Publication No. US20050250180A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M
; APPLICANT: LaVallie, Edward R
; APPLICANT: Racie, Lisa A
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 00766.000091.10
; CURRENT APPLICATION NUMBER: US/10/689,742
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: 09/746,783
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 50
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-689-742-50

Query Match      6.5%; Score 76.5; DB 1; Length 428;
Best Local Similarity 21.7%; Pred. No. 1.8;
Matches 55; Conservative 28; Mismatches 82; Indels 89; Gaps 14;

QY 3 LTPKEQEKFLYYAGEVARKKAE--GLKLNQPEAIAYISAHIMDEARRGKKTVA----- 55
DB 186 LTP--NTPFLIDNDGNTIENTEKLQYSGERLYKFTVTAYDC-----GKRAADDAEV 234
QY 56 --QLMEBC----MHFLKKDEVMPGVGNMVPDLGVVEATFPDGTKLVTNN---WPIEDPHF 106
```

Db 235 EIQVFTCPKPSQGNKRLEYAPGAGSL-----ALFP-GIRLETCDPLNLIQATIEL 286
QY 107 KAGEVKGCDKD-----IELNAGEKVELE-VTNEGPK-----SLHV---GSHFHF 149
Db 287 QTSYVAKGCDRDYSEARLKLCGAATGEVDLLPMEGPNANWTAGLSVHYSDSSLIYWF 346
QY 150 EANKALKFDEKAYKRLDIPSGNTLRIGAGQTRKVLIPGSKKVGNGLVNNADE 209
Db 347 NGTQAVQ-----VPLGGP-----SGLSGPQDS 369
QY 210 RHKKALDKAKSHG 223
Db 370 LSDHFTLSFWMKHG 383

RESULT 6
US-10-821-234-1089
; Sequence 1089, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes Version 1.0
; SEQ ID NO 1089
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1089

Query Match 6.2%; Score 73.5; DB 1; Length 449;
Best Local Similarity 21.2%; Pred. No. 3.8; Indels 61; Gaps 9;
Matches 51; Conservative 35; Mismatches 93
QY 1 VKLTPKEQKFLLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAOLMEE 60
Db 142 VKEPTQKQRLHVEQELTTE-----VEKIKTVKESATSEKLTVPVLLAQ 198
QY 61 CMHFLKKDEMPGVGNMV-PDLGVEATFPDGTGLVTNNWPIBDEHFKAKEVKGCDKDI 119
Db 189 -LAALQQLVASHLEKLLGPDAAINLTDPGALAKRLLQLLEATKNSKG----- 237
QY 120 ELNAGEKVELEVTNEGPKSLVGHSHFH-----PFEANKALKFDR---EKAYGKRLDI 169
Db 238 --SGGK-----TTGTPPDSLSLVTYELHSRPEQDKFSQAQVAELBKRLTELETAVRCQ 289
QY 170 PSGNTLRIGAGQTRKVLIPGSKK-----KVIGMGNLVNNAIDERHKKHA 215
Db 290 DQANPLUSAGLQACLMETVELLOAKVSLDALDVLQDVEARLQSVLGKVNIEA----KHKA 345

RESULT 7
US-10-858-730-6
; Sequence 6, Application US/10858730
; Publication No. US20050255568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.

; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Shewanella oneidensis
US-10-858-730-6

Query Match 6.1%; Score 72; DB 1; Length 418;
Best Local Similarity 21.7%; Pred. No. 4.7;
Matches 50; Conservative 34; Mismatches 82; Indels 64; Gaps 12;
QY 27 GLKLNQPEAIAY-----ISAHIMDEARRGKKTVAOLMEECMHFLLKDKDEMPGVGNMVPDL 81
Db 216 GAKVLHPDPSVEYAOQRFKVPURVLSSFEAGQOTLLQFGDESELAASAASVQGIANKALATL 275
QY 82 GVEATFPDGTGL-----VTNVN--PIEPDE-----HFKAKEVKGCDKDI 121
Db 276 TIEGLFTSSERYQALLACLARLEVDVFITPLKLEISPVESVFMLEAKV-----DILL 331
QY 122 NAGKEVTELEVTNEGPKSLVGHSHFHFEANKALKFDEKA-----YGRKLDIPSG---NT 174
Db 332 H-----EVLSE---SLDLGQ-----LIVERQRAKVSIVGKGLQAKVGLLTAKM 372
QY 175 LRIGAGQTRKVLIPGSKKVGNGLVNNAIDERHKKHAKDKAKSHGF 224
Db 373 LDVLGNETIHAKLSTSEK-----LSTVIDERDLHKAV-RALHHAF 413

RESULT 8
US-10-793-626-1248
; Sequence 1248, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1248
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1248

Query Match 6.0%; Score 71; DB 1; Length 311;
Best Local Similarity 18.9%; Pred. No. 3.9;
Matches 58; Conservative 35; Mismatches 100; Indels 114; Gaps 13;
QY 1 VKLTPKEQKFLLYYAGEVARKKAEGLKLNQPEAI--AYISAHIMDEARR----- 49
Db 6 LKNLPSEESYFLFY-----INRPSIIVGNKNQNTIEVNCQAYIDKKQIDVVRISGGGAVYH 61
QY 50 --GKKTVAOLMEECMHFLLK--KDEVMFPGVGNMVPDLGVEATFP-----DGTKL 93
Db 62 DTGNLNFSPITDDDGHSFHNFKFTMP-IVQALQSMGVNAEMWTGRNDIQVGQAKISGNAM 120

QY 94 VTVNWPIDPEHFKAGEVVKFGCD-----KDIELNAGK----- 125
Db 121 VKVK-----NRMFSGHTLMLNCDLNEVQKALKVNPDAKIKSGVKSVKRKVANIEBFLQEP 175
QY 126 -----EVTELEVTNPGPKSLHVGSHFHP-----FEANKALKFDRK 161
Db 176 IDIEBKILKTIIGENEVEEYIITEEDWKNIKQLSDKEYRTWENYGSNPKYNIEREE 235
QY 162 AYGK-----RLDIPSGNTLR-----ICAGQTRKVQLIPLGSKKVKVIGMGLVNNIADE 209
Db 236 KEKGFQIKLDVKKGRIERAKLFGDFGEGDVTELE-----HALVGLCHDF 282
QY 210 RHKKAL 216
Db 283 EHIEEAL 289

RESULT 9

US-10-793-626-1460
; Sequence 1460, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1460
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-1460

Query Match 5.9%; Score 70; DB 1; Length 329;
Best Local Similarity 25.2%; Pred. No. 5.3;
Matches 52; Conservative 19; Mismatches 89; Indels 46; Gaps 10;
QY 42 HIMDEARRGKKTVAQLMEECMHFLKDEVMVPGVGNMVPDLGVEATFPDGT-----KLVTYN 97
Db 85 YIKDITRRGKVPi-----IAGGTGLYIQSLLYNAYFEDESISEDKMKQVK 129
QY 98 WPIEPEHP---KAGEVKFGCD---KDIELNAGKEVTE-LEVTNEGPKSLHVGSHFHPF 149
Db 130 LXLKELEHNNKHLHEYLASFKESAKDHPNNRKEVLRRAIEYLYTKKLLSRKKVQQP 189
QY 150 EANK-----ALKFDRKAY---GKRLDIPSGNTLRIGAGQTRKVQ-LIPLG-----GSKKV 196
Db 190 TENYDTLLIGIEMSRETLYLRINKRVDI-----MLGHGLENEVQHLVQGGPEASQMQA 243
QY 197 IGWNGLVNNIADIERHKHAKLDIAKSH 222
Db 244 IGYKELVPVIGKNISMENAVEKLQKH 269

RESULT 10

US-10-821-234-1608
; Sequence 1608, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A

; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PRT_SEQ_genes Version 1.0
; SEQ ID NO 1608
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1608

Query Match 5.9%; Score 70; DB 1; Length 437;
Best Local Similarity 20.5%; Pred. No. 7.8;
Matches 46; Conservative 31; Mismatches 65; Indels 82; Gaps 11;
QY 4 TPKEQEKELLYAGEVARKKAEGLKLNQPEATAYISAHIMDEARRGK--KTVAQLMEEC 61
Db 46 TPEFLRRK---PAGKVPAFEGDDGFCVFESNAIAY---YVSNEELRGSTPEAAQV--QW 98
QY 62 MHFLKKDEVMPGVGNMVPDLGV-----EAT 86
Db 99 VSFADSDIVPPASTWVPPTLGIHNNKQATENAKEEVERILGLLDAYLKTFTFLVGERVT 158
QY 87 FPDGTKLVTYNW----PIEP-----DEHFKA--GEVFGCKDKDIELN 122
Db 159 LADITVCTLLWLYKQVLEPSFQAFTNTRWELTCINQPFRAVLGEVKL-CEKMAQFD 217
QY 123 AGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKFDRKAYGKR 166
Db 218 A-KKFAETOPKDTPRK-----EKSREEKQKPOAER 248

RESULT 11

US-10-793-626-244
; Sequence 244, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 244
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-244

Query Match 5.8%; Score 69; DB 1; Length 501;
Best Local Similarity 20.0%; Pred. No. 12;
Matches 42; Conservative 38; Mismatches 74; Indels 56; Gaps 9;
QY 30 LNOPEAIAYISAHIMDEARRGKKTVAQL-----MEECMHFLKKDEVMPG 73
Db 304 VTEPDEIRDSLAIVLDESRLNRLVNNELNVARMDAAGLSVEKELOPQHLLDKMESKYR 363
QY 74 VGNMVPDLGVEATFP-----DGTKLVTNNWPIEPEHFKAGE-VKFGCKDIE 120
Db 364 MQS--BELGLTMTFDSNNDEQLWNYDMRMDQVLTNLDINATRYTQAGDSIKISIDSD 421
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHF-HFPEANKALKFDR-----KAYGKR 166
Db 422 FN-----ILTTDTGTGTGIAPEHLKQVDFRKYVKAARKGKGGTGLGLFICKMIIEHGG 477
QY 167 LDIPSGNLTIGAGQTRKVQLIPLGGSKKV 196

QY	107	KAGEVFGCDKDIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFPEANKALKFDREK 16	107	KAGEVFGCDKDIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFPEANKALKFDREK 16
DB	628	KDGTVTIIFGLSQDSGLTIGKSTLNNDGLTVKDTNE---QIQVGAN-----GIKF--TN 674	628	KDGTVTIIFGLSQDSGLTIGKSTLNNDGLTVKDTNE---QIQVGAN-----GIKF--TN 674
QY	162	AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG---GSKK 195	162	AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG---GSKK 195
DB	675	VNGSNPCTGANTYARITRDKIGFAGSDGAVDTNPKYLDQDKLVGNVKITNTGINAGGKA 734	675	VNGSNPCTGANTYARITRDKIGFAGSDGAVDTNPKYLDQDKLVGNVKITNTGINAGGKA 734
QY	196	VIGMGLVNNIADERHKKHAL-----DKAKSH 222	196	VIGMGLVNNIADERHKKHAL-----DKAKSH 222
DB	735	ITGLSPTLPISADOSSRNIELGNTIQDKKSN 766	735	ITGLSPTLPISADOSSRNIELGNTIQDKKSN 766

RESULT 14
 US-11-013-759--4
 ; Sequence 4, Application US/11013759
 ; Publication No. US20050249747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Sasaki, Ken
 ; APPLICANT: Yang, Yan Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
 ; TITLE OF INVENTION: PROTEIN OF MORAXELLA
 ; FILE REFERENCE: 1038-921MIS:jb
 ; CURRENT APPLICATION NUMBER: US/11/013,759
 ; CURRENT FILING DATE: 2004-12-16
 ; PRIOR APPLICATION NUMBER: US/09/361,619
 ; PRIOR FILING DATE: 1999-07-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2047
 ; TYPE: PRT
 ; ORGANISM: Moraxella catarrhalis
 US-11-013-759--4

QY	107	KAGEVFGCDKDIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFPEANKALKFDREK 161	107	KAGEVFGCDKDIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFPEANKALKFDREK 161
DB	683	KDGTVTIIFGLSQDSGLTIGKSTLNNDGLTVKDTNE---QIQVGAN-----GIKF--TN 729	683	KDGTVTIIFGLSQDSGLTIGKSTLNNDGLTVKDTNE---QIQVGAN-----GIKF--TN 729
QY	162	AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG---GSKK 195	162	AYGKRLDIPSGNTLRI-----CAGOTRK-----VOLIPLG---GSKK 195
DB	730	VNGSNPCTGANTYARITRDKIGFAGSDGAVDTNPKYLDQDKLVGNVKITNTGINAGGKA 789	730	VNGSNPCTGANTYARITRDKIGFAGSDGAVDTNPKYLDQDKLVGNVKITNTGINAGGKA 789
QY	196	VIGMGLVNNIADERHKKHAL-----DKAKSH 222	196	VIGMGLVNNIADERHKKHAL-----DKAKSH 222
DB	790	ITGLSPTLPISADOSSRNIELGNTIQDKKSN 821	790	ITGLSPTLPISADOSSRNIELGNTIQDKKSN 821

RESULT 15
 US-11-013-759--7
 ; Sequence 7, Application US/11013759
 ; Publication No. US20050249747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Loosmore, Sheena M.
 ; APPLICANT: Sasaki, Ken
 ; APPLICANT: Yang, Yan Ping
 ; APPLICANT: Klein, Michel H.
 ; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
 ; TITLE OF INVENTION: PROTEIN OF MORAXELLA
 ; FILE REFERENCE: 1038-921MIS:jb
 ; CURRENT APPLICATION NUMBER: US/11/013,759
 ; CURRENT FILING DATE: 2004-12-16
 ; PRIOR APPLICATION NUMBER: US/09/361,619
 ; PRIOR FILING DATE: 1999-07-27
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7

```
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-7

Query Match      5.8%; Score 68.5; DB 7; Length 2047;
Best Local Similarity 25.0%; Pred. No. 96;
Matches 38; Conservative 18; Mismatches 47; Indels 49; Gaps 8;

QY 107 KAGEVKFGCDKIELNAGKEVTE-----LEVTNEGPKSLHVGSHPHFFFEANKALKFEDREK 161
Db   ||| ||| :| ||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
683 KGGTWTFTGLSQDSGLTIGKSTLNNDGLTVKDTNE---QIQVGAN-----GIKF--TN 729

QY 162 AYGRKLDIPSGNTLRI-----GAGQTRK-----VQLIPLG---GSKK 195
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
730 VNGSNPGTGIANTRITRDKIGFAGSDGAVDTNKPYPDQDKLQVGNVKITNTGINAGGKA 789

QY 196 VIGMGLVNNIADERHKHKAL-----DKAKSH 222
Db   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
790 ITGLSPTLPSIADQSSRNIELGNTIQDKDKSN 821
```

Search completed: November 28, 2005, 08:25:57
Job time : 3.84635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:13:38 ; Search time 18.5013 Seconds
(without alignments)
1009.915 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADERHKHKKALDKAKSHGFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/aaa/5 COMB.pep.*
- 2: /cgn2_6/prodata/1/aaa/6 COMB.pep.*
- 3: /cgn2_6/prodata/1/aaa/H COMB.pep.*
- 4: /cgn2_6/prodata/1/aaa/ECTUS_COMB.pep.*
- 5: /cgn2_6/prodata/1/aaa/RE_COMB.pep.*
- 6: /cgn2_6/prodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	622	52.7	234	2	US-09-338-920B-11
2	610.5	51.7	238	1	US-08-920-095-2
3	610.5	51.7	238	2	US-09-338-920B-5
4	610.5	51.7	238	4	PCT-US96-05800-2
5	610.5	51.7	245	2	US-09-431-705-4
6	603.5	51.1	238	2	US-09-338-920B-7
7	600	50.8	237	1	US-08-467-822-22
8	600	50.8	237	2	US-08-432-697-22
9	600	50.8	237	2	US-08-466-248-22
10	573.5	48.6	228	2	US-10-012-819-154
11	572.5	48.5	237	1	US-08-467-822-20
12	572.5	48.5	237	2	US-08-432-697-20
13	572.5	48.5	237	2	US-08-466-248-20
14	566.5	48.0	237	2	US-09-338-920B-9
15	544.5	46.1	213	2	US-09-338-920B-13
16	431.5	36.5	840	1	US-08-467-822-25
17	431.5	36.5	840	2	US-08-432-697-25
18	431.5	36.5	840	2	US-08-466-248-25
19	281.5	23.8	100	1	US-08-967-513-3
20	281.5	23.8	100	1	US-08-687-645B-3
21	281.5	23.8	103	2	US-09-489-039A-9261
22	269.5	22.8	99	1	US-07-732-242C-1
23	266.5	22.6	107	2	US-09-543-681A-6101
24	264.5	22.4	132	2	US-09-134-001C-5049
25	261.5	22.1	100	1	US-08-467-822-23
26	261.5	22.1	100	2	US-08-432-697-23
27	261.5	22.1	100	2	US-08-466-248-23

28	261	22.1	145	2	US-09-134-001C-4982	Sequence 4982, Ap
29	255.5	21.6	103	2	US-09-328-352-5887	Sequence 5887, Ap
30	250.5	21.2	100	2	US-09-602-777A-18	Sequence 18, Appl
31	250.5	21.2	125	2	US-09-252-991A-26884	Sequence 26884, A
32	247	20.9	111	2	US-09-710-279-116	Sequence 116, App
33	225.5	19.1	153	2	US-09-543-681A-6085	Sequence 6085, Ap
34	224	19.0	106	1	US-07-732-242C-2	Sequence 2, Appli
35	224	19.0	119	2	US-09-328-352-5915	Sequence 5915, Ap
36	212.5	18.0	162	2	US-09-602-777A-10	Sequence 10, Appl
37	209	17.7	106	1	US-08-967-513-4	Sequence 4, Appli
38	209	17.7	106	1	US-08-687-645B-4	Sequence 4, Appli
39	208	17.6	109	2	US-09-489-039A-9266	Sequence 9266, Ap
40	206	17.4	109	1	US-08-467-822-24	Sequence 24, Appl
41	206	17.4	109	2	US-08-432-697-24	Sequence 24, Appl
42	206	17.4	109	2	US-08-466-248-24	Sequence 24, Appl
43	185	15.7	137	2	US-09-252-991A-26886	Sequence 26886, A
44	92	7.8	289	2	US-09-134-001C-4415	Sequence 4415, Ap
45	90.5	7.7	790	1	US-08-363-560-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-338-920B-11
; Sequence 11, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-11

Query Match	52.7%	Score 622;	DB 2;	Length 234;
Best Local Similarity	54.7%;	Pred. No. 2.9e-62;		
Matches 123;	Conservative 41;	Mismatches 59;	Indels 2;	Gaps 2;
QY	1	VKLTPKEQKFLLYAGEVARKKAEGLKLNQPEAIAVISAHIMDEARRGKKTVAQLMEE	60	
DB	1	MKUTPKELDKMLHYAGELAKQKAKGKLNYYTAVALISAHVWEERAGKKSVDLQWE	60	
QY	61	CMHFLKKDEYMPGVGNMVPDLGVEATFPDGTGKLVTVNMPPEP-DEHFKAGEYKFGCDKI	119	
DB	61	GRILLKADDVMPGVVAHMHVEVIEAGPDGTGKLVTIHTPFVEAGSKLAPGEVILK-NEDI	119	
QY	120	ELNAGKEVTELVNTEGPKSLHVGSHFFHFFANKALPDRKAYGKRLDIPSGNTILRGA	179	
DB	120	TLNAGKHAVLKVKVKNKGDPRPVQVSGSHFFFEVFNKLLDPDRKAYGKRLDIASGTAVREP	179	
QY	180	GQTRKVLQILPLGSKKVIYGNGLVNNIADERHKHKKALDKAKSHGF	224	
DB	180	GEEKTVELIDGNGKRIYGFNALVDROADHGGKKLALRAKEKH	224	

RESULT 2

US-08-920-095-2
; Sequence 2, Application US/08920095
; Patent No. 5837240
; GENERAL INFORMATION:
; APPLICANT: Cynthia K. Lee et al.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE

```
;
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,095
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,041
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-920-095-2

Query Match      51.7%; Score 610.5; DB 1; Length 238;
Best Local Similarity 52.7%; Pred. No. 6.1e-61;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLLYAGAEVARKKAEGKLKNOPEAIYSAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKKKEGKIKLNYVEAVALISAHIMEERAGKKTAAELMQE 60
QY 61 CMHFLKDEYMPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDHFHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFKANKALFKDREKAYGKRLDIPSGNTLRIAG 180
Db 121 INEGKAVSVKVNVDGDPVQIGSHFFHFFVNRCLDFDREKTFGKRLDIASGTAVRFEPG 179
QY 181 QTRKVLQIPLGSKKVGNGLVNNTADERHKKHAKDKAKSHGF 224
Db 180 EKSVELIDIGGNRRIFGNALVDRQADNESKKIALHRAKERGF 223

RESULT 4
PCT-US96-05800-2
; Sequence 2, Application PC/TUS9605800
; GENERAL INFORMATION:
; APPLICANT: Oravax, Inc.
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05800
; FILING DATE: 23-APR-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,041
; FILING DATE: 28-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/568,122
; FILING DATE: 06-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06132/020001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
PCT-US96-05800-2

Query Match      51.7%; Score 610.5; DB 4; Length 238;
Best Local Similarity 52.7%; Pred. No. 6.1e-61;
Matches 118; Conservative 43; Mismatches 62; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLLYAGAEVARKKAEGKLKNOPEAIYSAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDKMLHYAGELAKKKEGKIKLNYVEAVALISAHIMEERAGKKTAAELMQE 60
QY 61 CMHFLKDEYMPGVGNMVPDLGVEATFPDGTGKLVTVNWPIDEPDHFHFKAGEVKGCDKDIE 120
Db 61 GRTLLKPDVMDGVASMIHEVGIEAMFPDGTGKLVTVHTPIEANGKLVPGEL-FLKNEDIT 119
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFFHFFKANKALFKDREKAYGKRLDIPSGNTLRIAG 180
Db 121 INEGKAVSVKVNVDGDPVQIGSHFFHFFVNRCLDFDREKTFGKRLDIASGTAVRFEPG 179
QY 181 QTRKVLQIPLGSKKVGNGLVNNTADERHKKHAKDKAKSHGF 224
Db 180 EKSVELIDIGGNRRIFGNALVDRQADNESKKIALHRAKERGF 223

RESULT 3
US-09-338-920B-5
; Sequence 5, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
```

```

; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-09-338-920B-7

Query Match      51.1%; Score 603.5; DB 2; Length 238;
Best Local Similarity 52.2%; Pred. No. 3.8e-60;
Matches 117; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

```

```

Db      1  MKUTPKELDKLMLHYAGELAKKREKGKLNVEAVRLISAHIMEAREGKKTAAELMQE  60
Qy      61  CMEHFKDEYMPGVGNVMPDVGBEATPPDGTKLVTNNMPIEPDEHFKAAGEVKGCDKDIE  120
Db      61  GRTLLKPPDDVMDGVASMIHEVGIEAMPDGTKLVTHTPIEANGKLVPGEL-FLKNEDIT  119
Qy      121  LNAKGKVELEVTNEGPKSLHVGHSHFFHFEANKALKDFREKAYGRKLDIPSGNTLIGAG  180
Db      120  INEGKKAHSVKKVPVGDPRPQIGSHGHHFFEVNRCLDFFREKTFGKRDLIASGTAVRPEFG  179
Qy      181  QTRKVQLIPGGSKVKVGMGNLVNNIADERHKHKALDKAKSHGF  224
Db      180  EEKSVELIDIGGNRRRIFGFNALVDRQDANESKKTALHRAKERGF  223

RESULT 7
US-08-467-822-22
; Sequence 22, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.

```

TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 TREATMENT OF INFECTIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 POLYPEPTIDES
 TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 POLYPEPTIDES
 NUMBER OF SEQUENCES: 44
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ADDRESSEE: Dunner
 STREET: 1300 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995

```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-822-22

Query Match      50.8%; Score 600; DB 1; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPEQEKFLLYAGGEVARKKAEGLKLNQPEALAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDK-LMHYAGELAKRKEGKIKLVYEAVALISAHIMEERAGKKTAAELMQE 59
QY 61 CMHFLKDEVPVGVMVDPDLGVEATFPDGTKLVTVMPIEPDHFHFKAGEVKGCDKDIE 120
Db 60 GRTLLKPDVDVMDGVASMIHEVGIEAMPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 118
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFKDREKAYGKRLDIPSGNTLRIGAG 180
Db 119 INEGKAVSVKVNVDGPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFEPG 178
QY 181 QTRKVLQIPLGSKKVIQGMGLVNNIADERHKKHAKDKAKSHGF 224
Db 179 EEKVELIDIGGNRRIFGFNALVDROADNESKKIALHRAKERGF 222

RESULT 8
US-08-432-697-22
; Sequence 22, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-822-22

Query Match      50.8%; Score 600; DB 1; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPEQEKFLLYAGGEVARKKAEGLKLNQPEALAIYISAHIMDEARRGKKTVAQLMEE 60
Db 1 MKLTPKELDK-LMHYAGELAKRKEGKIKLVYEAVALISAHIMEERAGKKTAAELMQE 59
QY 61 CMHFLKDEVPVGVMVDPDLGVEATFPDGTKLVTVMPIEPDHFHFKAGEVKGCDKDIE 120
Db 60 GRTLLKPDVDVMDGVASMIHEVGIEAMPDGTKLVTVHTPIEANGKLVPGEL-FLKNEDIT 118
QY 121 LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALFKDREKAYGKRLDIPSGNTLRIGAG 180
Db 119 INEGKAVSVKVNVDGPRVQIGSHFHFPEVNRCLDFDREKTFGKRLDIASGTAVRFEPG 178
QY 181 QTRKVLQIPLGSKKVIQGMGLVNNIADERHKKHAKDKAKSHGF 224
Db 179 EEKVELIDIGGNRRIFGFNALVDROADNESKKIALHRAKERGF 222

RESULT 9
US-08-466-248-22
; Sequence 22, Application US/08466248
; Patent No. 6258359
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-466-248-22

Query Match 50.8%; Score 600; DB 2; Length 237;
Best Local Similarity 52.7%; Pred. No. 9.4e-60;
Matches 118; Conservative 43; Mismatches 61; Indels 2; Gaps 2;

QY 1 VKLTPKEQKFLLYAGAEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDK-LMHYAGELAKKKEGKIKLNYVEAVALISAHIMEEAEARAGKKTAAELMQE 59

QY 61 CMHFLKKDEYMPGVGNMVPDLGVETFPDGTGLVTWNPIEPDEHFKAGVKGCDKIDIE 120
DB 60 GRTLLKPDVDMGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDIT 118

QY 121 LNAGKEVTELVNTEGPKSLHVGSHFPPFANKALKPDKREKAYGKRLDIPSGNTLRIGAG 180
DB 119 INEGKAVSVKXNVGDRPVQIGSHFHFVNRCLDFDREKTKRDLIASGTAVRPEPG 178

QY 181 QREKVQLIPLGSKKVGWGLNNTADERHKHKALDKAKSHGF 224
DB 179 EKSVELIDIGNRRIFGNALVDRQADNESKKIALHRAKERGF 222

RESULT 10
US-10-012-819-154
Sequence 154, Application US/10012819
Patent No. 6916615
GENERAL INFORMATION:
APPLICANT: Legrain, Pierre
APPLICANT: Selig, Luc
TITLE OF INVENTION: Collection of Prokaryotic DNA for Two-Hybrid System, Helicobacter
TITLE OF INVENTION: pylori Protein-Protein Interactions and Applications thereof
FILE REFERENCE: B5053
CURRENT APPLICATION NUMBER: US/10/012,819
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: EP 99401066.8
NUMBER OF SEQ ID NOS: 278
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 154
LENGTH: 228
TYPE: PRT
ORGANISM: Helicobacter pylori
US-10-012-819-154

Query Match 48.6%; Score 573.5; DB 2; Length 228;
Best Local Similarity 52.1%; Pred. No. 9.2e-57;
Matches 111; Conservative 41; Mismatches 60; Indels 1; Gaps 1;

QY 12 LLYYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEECMHFLKDEVM 71
DB 2 MLHYAGELAKKKEGKIKLNYVEAVALISAHIMEEAEARAGKKTAAELMQEGRTLLKPDVM 61

QY 72 PGVGNMVPDLGVETFPDGTGLVTWNPIEPDEHFKAGVKGCDKIDIELNAGKEVTELE 131
DB 62 DGVASMIHEVGIEAMPDGTGLVTHTPIEANGKLVPGEL-FLKNEDITINEGKAVSVK 120

QY 132 VTNEGPKSLHVGSHFHFVNRCLDFDREKAYGKRLDIPSGNTLRIGAGQTRKVLPIGL 191

DB 121 VKNVGDRPVQIGSHFHFVNRCLDFDREKTKRDLIASGTAVRPEPCEEKSVELIDIG 180

QY 192 GSKKVIQMGWGLNNTADERHKHKALDKAKSHGF 224
DB 181 GNRIFGNALVDRQADNESKKIALHRAKERGF 213

RESULT 11
US-08-467-822-20
Sequence 20, Application US/08467822
Patent No. 5843460
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,822
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..237
OTHER INFORMATION: /note= "URE A - FIGURE 3."
US-08-467-822-20

Query Match 48.5%; Score 572.5; DB 1; Length 237;
Best Local Similarity 50.4%; Pred. No. 1.3e-56;
Matches 113; Conservative 43; Mismatches 67; Indels 1; Gaps 1;

QY 1 VKLTPKEQKFLLYAGAEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQLMEE 60
DB 1 MKLTPKELDKMLHYAGELAEERLARGVKLNYTEAVALISGRVMEKARDGNKSVADLMQE 60

Query Match 46.1%; Score 544.5; DB 2; Length 213;
Best Local Similarity 52.3%; Pred. No. 1.6e-53;
Matches 112; Conservative 36; Mismatches 65; Indels 1; Gaps 1;

Qy 13 LYTAGEVANRKRGAEGKLNQPEAIVISAHIMDEARRGKTVAQLMEECHFLKKDEVMP 72
:
: :

Search completed: November 28, 2005, 08:25:40
Job time : 19.5013 secs

Query Match	48.08;	Score 566.5;	DB 2;	Length 237;
Best Local Similarity	50.08;	Pred. No. 6.1e-56;		
Matches 112;	Conservative 43;	Mismatches 68;	Indels 1;	Gaps 1;
Qy	1	VKLTPEQEKFLIYAYAGEVARKKAEGLKLNQPEAIAYISAHIMDEARRGKKTVAQIMLEE	60	
Db	1	MKLTPEKLDKMLHYAGRLAEALARGVKLNYTEAVALISGRVMEKARDGNKSVADLMQE	60	
Qy	61	CMHFLKKDEYMPGVGNVMPDOLGVEATPPDGTGLVTVNWPIEPDQBHFKAGEVKGCGKDIE	120	
Db	61	GRTWLKKEENWMDGVASMIHEVGVIEANPPDGTGLVTIHTPVEDNGKLAPGEV-FLKNEDIT	119	
Qy	121	LNAGKEVTELEVNTNEGPKSLHVSGSHFFHFANAKLKFDRKAYGKRLDIPSGNTLRIAG	180	
Db	120	INAGKEAISLVKNKGDRPVQVGSFHFFHFVFNKLLDFDRAKSFCKRLDIASGTAVREPG	179	
Qy	181	QTRKVQLIPLGGSKKVTGMNGLVNNIADERHGHKALDKAKSHGF	224	
Db	180	EEKSVELIDIGNKRIYGFNSLVYDROADADGKGLKRAKEKF	223	

RESULT 15
US-09-338-920B-13
; Sequence 13, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Sonan, Gopalan
; APPLICANT: Sonan, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 061132/023002

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:00:02 ; Search time 36.4836 Seconds
(without alignments)
1497.963 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297	76.6	569	1 URKCBP	urease (EC 3.5.1.5
2	2244	74.8	569	1 S35291	urease (EC 3.5.1.5
3	2028.5	67.6	569	2 C36950	urease (EC 3.5.1.5
4	2016.5	67.2	571	2 F83681	urease alpha subu
5	1981.5	66.1	569	2 D69729	urease (alpha subu
6	1927.5	64.3	569	2 S75169	urease (EC 3.5.1.5
7	1897	63.3	568	2 AG2264	urease alpha chain
8	1889.5	63.0	569	2 S47104	urease (EC 3.5.1.5
9	1832	61.1	566	2 H83037	urease alpha subu
10	1832	61.1	568	2 C75586	urease, alpha subu
11	1826	60.9	572	2 H64075	urease (EC 3.5.1.5
12	1825	60.9	567	1 C36138	urease (EC 3.5.1.5
13	1822	60.8	570	2 S42607	urease (EC 3.5.1.5
14	1813	60.5	567	1 D43719	urease (EC 3.5.1.5
15	1809	60.3	567	1 S08480	urease (EC 3.5.1.5
16	1809	60.3	568	2 G85654	probable urease st
17	1809	60.3	568	2 D90794	urease alpha subu
18	1799.5	60.0	838	2 A96699	probable urease F1
19	1792.5	59.8	569	2 A97648	urease alpha chain
20	1792.5	59.8	569	2 AG2871	urease alpha subu
21	1785	59.5	570	2 AF3458	urease (EC 3.5.1.5
22	1784.5	59.5	571	2 E90027	urease alpha subu
23	1780.5	59.4	571	2 S38485	urease (EC 3.5.1.5
24	1779.5	59.3	572	2 AC0325	urease (EC 3.5.1.5
25	1753.5	58.5	840	1 URJB	urease (EC 3.5.1.5
26	1715	57.2	573	2 A13332	urease (EC 3.5.1.5
27	1710.5	57.0	835	2 T37939	urease (EC 3.5.1.5
28	1708	57.0	573	2 S36028	urease (EC 3.5.1.5
29	1696.5	56.6	598	2 G82890	urease complex com

RESULT 1

URKCBP

urease (EC 3.5.1.5) 62K chain - Helicobacter pylori (strains 26695, J99, and others)
N;Alternate names: urease beta chain; urease chain B; urease large subunit
C;Species: Helicobacter pylori
C;Date: 30-Sep-1991 #sequence revision 02-Dec-1994 #text change 05-Oct-2004
C;Accession: B38537; H64528; A71977; F41834; S07885; S12487; A49215; B61371; A35306; B41

R;Labigne, A.; Cussac, V.; Courcoux, P.

J. Bacteriol. 173, 1920-1931, 1991

A;Title: Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsib

A;Reference number: A38537; MUID:91161505; PMID:2001995

A;Accession: B38537

A;Molecule type: DNA

A;Residues: 1-569 <LAB>

A;Cross-references: UNIPROT:P14917; UNIPARC:UPI000002C20A; GB:M60398; NID:g149007; PIDN:

R;Tomb, J.P.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A;Reference number: A64520; MUID:97394467; PMID:9252185

A;Accession: H64528

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-569 <TOM>

A;Cross-references: UNIPARC:UPI000002C20A; GB:AE000511; NID:g2313152; PIDN:

A;Experimental source: strain 26695

R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A;Reference number: A71800; MUID:99120557; PMID:9923682

A;Accession: A71977

A;Molecule type: DNA

A;Residues: 1-569 <ARN>

A;Cross-references: UNIPARC:UPI000002C20A; GB:AE001446; GB:AE001439; NID:g4154573; PIDN:

A;Experimental source: strain J99

R;Cussac, V.; Ferrero, R.L.; Labigne, A.

J. Bacteriol. 174, 2466-2473, 1992

A;Title: Expression of Helicobacter pylori urease genes in Escherichia coli grown under

A;Reference number: A41834; MUID:92210488; PMID:1313413

A;Accession: F41834

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 565-569 <CUS>

A;Cross-references: UNIPARC:UPI000002D2AD; GB:M84338

R;Clayton, C.L.; Pallen, M.J.; Kleanthous, H.; Wren, B.W.; Tabaqchali, S.

Nucleic Acids Res. 18, 362, 1990

A;Title: Nucleotide sequence of two genes from Helicobacter pylori encoding for urease

A;Reference number: S07884; MUID:90221820; PMID:2326167

A;Accession: S07885

Db	123	AEGLIVTAGGIDTHIHIFISPOQIPTAFASGVTTWIMGGTGPADGTNATTITPGRANLKSM	182
Qy	182	LRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEYD	241
Db	183	LRAAEYAMNLGLFAGKGNVSYBPSLRDIEAGAIGFKIHEDWGSTPAAIHHCLNVADEYD	242
Qy	242	VQVCIHTDVTNEAGYVDDTLNMMNGRAIHAYHIEGAGGHSPPDVITWAGELNLPSTTP	301
Db	243	VQVAIHTDVTNEAGVVEDTLTAAGRTIHTFHTEGAGGSHAPDVIKMGSEFNILPASTNP	302
Qy	302	TIPYTINTVAEHLDMLTCHHLDKRIREDLFOSOSRIRPGSIAAEDVLHDMGVAMTSSD	361
Db	303	TIPFTKNTAEHMDMLMVCCHLHDKSIEDVQFADSRIRPOTIAAEQDLHDMGIFSISSD	362
Qy	362	SOAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY	421
Db	363	SOAMGRVGEVITRTWQTADKNKKEFGKLEKEGDNDNFRIKRYISKYITINPGIAHGISDY	422
Qy	422	IGSVEEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYYREMPFGHH	481
Db	423	VGSVEGVKVALWSPAPFGIKPKPMIIGKGFIALSQMGDANASIPTPQPVYYREMPFGHH	482
Qy	482	GKAKFDTSTFVSKVAYENGVEKGLERQVLVKNCRNITKKPKFNDKTAKITVDPKPT	541
Db	483	GKNKFDNTITFVSQAAYKAGIKBELGLDRAAPPVKNCRNITKKDLKPNVDVTAHIDVNPET	542
Qy	542	FEVFDGKCLTSKPTSOVPLAQRITFF	568
Db	543	YKVKVDGKEVTSKADELSLAQLYNLF	569
RESULT 3			
C36950			
N:Alternate names: ureC protein			
C:Species: Bacillus sp.			
C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 05-Oct-2004			
C:Accession: C36950			
R:Maeda, M.; Hidaka, M.; Nakamura, A.; Masaki, H.; Uozumi, T.			
J. Bacteriol. 176, 432-442, 1994			
A>Title: Cloning, sequencing, and expression of thermophilic Bacillus sp. strain TB-90 u			
A:Reference number: A36950; MUID:9411379; PMID:8288539			
A:Accession: C36950			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-569 <MAE>			
A:Cross-references: UNIPROT:Q07397; UNIPARC:UPI0000137D67; GB:D14339; NID:g393296; PIDN:			
C:Superfamily: urease, alpha subunit; urease 62K chain homology			
C:Keywords: hydrolase			
P:5-552/Domain: urease 62K chain homology <u62>			
Query Match 67.6%; Score 2028.5; DB 2; Length 569;			
Best Local Similarity 66.1%; Pred. No. 2.2e-137;			
Matches 375; Conservative 77; Mismatches 112; Indels 3; Gaps 2;			
Qy	3	MKKQEVNTYPTKDGKVLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMGSN-SPD	61
Db	5	MRRQADYMPFPVGAIRLADSELFIETKDYTYGDEVKGGKVLDRGMGQHPATS	64
Qy	62	ENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVVGVTGTEALA	121
Db	65	DECDVLVLTNAIIVDTYGIYKADIGIKDGMISIGKAGNPLMDGVD--NVIGAAETVIA	122
Qy	122	GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTNATTITPGKWNLRH	181
Db	123	AEGMIITAGGIDAHIFICPQIETALASGVTIMIGGTTGATGTNATCTCPGFWNIHRM	182
Qy	182	LRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEYD	241
Db	183	LQAAEEFPINLGLFGKNCSDSEAPLEQIEQIEAGAVGLKLHEDWGSTAAAIIDTCLKVADRYD	242
Qy	242	VQVCIHTDVTNEAGYVDDTLNMMNGRAIHAYHIEGAGGHSPPDVITWAGELNLPSTTP	301

Db	243	VQVAIHTDVTNEGGFVEDTLKAIDGRVITHYHTEGAGGSHAPDIIKAAGFPNLPSSTNP	302
Qy	302	TIPYTINTVAEHLDMLTCHHLDKRIREDLFOSOSRIRPGSIAAEDVLHDMGVAMTSSD	361
Db	303	TRPYTINTLEHLDMLMVCCHLDANIPEDIAFADSRIRKETIAAEDVLHDLGVFWMISSD	362
Qy	362	SOAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKDNDFRIKRYISKYITINPALTHGVSEY	421
Db	363	SOAMGRVGEVITRTWQTADKNKQKGKLOEDNGVDNFRVKRYIAKYITINPAIAGIADY	422
Qy	422	IGSVEEGKIADLVVNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYYREMPFGHH	481
Db	423	VGSVEVEKGLADLVVNPAPFGVKPELVKGMIAYSTMGDPNASIPTPQPVLYRMPFAAK	482
Qy	482	GKAKFDTSTFVSKVAYENGVEKGLERQVLVKNCRNITKKPKFNDKTAKITVDPKPT	541
Db	483	GDAKYQTSITFVSKAAVEKEGHEQLGKKVKVPHGIRKLTCKDLILNDKTKPIDVDPQT	542
Qy	542	FEVFDGKCLTSKPTSOVPLAQRITFF	568
Db	543	YEKVDGQLVTCPEAETVPMAQRYFLF	569
RESULT 4			
P83681			
urease alpha subunit ureC [imported] - Bacillus halodurans (strain C-125)			
C:Species: Bacillus halodurans			
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 05-Oct-2004			
C:Accession: P83681			
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Hira			
Nucleic Acids Res. 28, 4317-4331, 2000			
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and			
A:Reference number: A83650; MUID:20512582; PMID:11058132			
A:Accession: P83681			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-571 <STO>			
A:Cross-references: UNIPARC:UPI000000C3801; GB:AP001507; GB:BA000004; NID:g10172612; PIDN:			
A:Experimental source: strain C-125			
C:Genetics:			
A:Gene: ureC			
C:Superfamily: urease, alpha subunit; urease 62K chain homology			
Query Match 67.2%; Score 2016.5; DB 2; Length 571;			
Best Local Similarity 64.4%; Pred. No. 1.6e-136;			
Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2;			
Qy	1	MKKKQEVNTYPTKDGKVLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMGS--N	58
Db	1	MKLTRAQHASLYGPTGVGKVLADTDLLEIEKDYTYGDECKGGKVLDRGMGQSAYV	60
Qy	59	SPENTLDLVITNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDDGVSPHMVVGVTGTE	118
Db	61	TRDEGVLDLITNATIIDYTGIVKADIGIKDGHIVGIGKGNPDIMDGVEHMIYGASTE	120
Qy	119	ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTNATTITPGKWNL	178
Db	121	AIAGEGLIVTAGGIDAHIFISPOQIDVAIASGVTIMIGGTTGATGTNATCTCTGKWN	180
Qy	179	HRMLRAAEYSMNVLGLKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD	238
Db	181	ERMLEAADAPVNLGFLGKGNASTPAPLREQIEAGAILGLHEDWGTTPAAIRTCLSVAD	240
Qy	239	EYDVQVCITDVTNEAGYVDDTLNMMNGRAIHAYHIEGAGGHSPPDVITWAGELNLPSS	298
Db	241	RMVQVAIHTDVTNEAGFVEDTIKAIGDRVITHYHTEGAGGSHAPDIMKVAGLPNLPSS	300
Qy	299	TTPTIPTINTVAEHLDMLTCHHLDKRIREDLFOSOSRIRPGSIAAEDVLHDMGVAMT	358
Db	301	TNTPRPTVNTIDSHLDMLMVCCHLDPNVPEDVAFADSRIRPETIAAEDIDQDLGVSM	360
Qy	359	SSDSQAMGRAGEVIPRTWQTADKNKKEFGKLPED--GKDNDFRIKRYISKYITINPALTHG	417

Db 361 SSDSQMGRVGEVIIRTWQTADMKKQKORALAEQQKGNNDNVRIRKYVSKYTIINPAITHG 420
Qy 418 VSEYIGSVEEGKIADLVVWNPAGFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREM 477
Db 421 IDYYGVSVEGKLADLVWDPFRFGVKPELILKGLIAYSQMGDPNASIPTPOPVSRRPM 480
Qy 478 FGHGKAKFDTSTITFVSKVAYENGVEKLGLEBRQVLVVKNCRNITKKDKFKNDKTAKITV 537
Db 481 FGSFGRARGTSTITFLSKAAMDVLGVEALGLQKIAHVRNCRSISKRSMKYNDATPNIEI 540
Qy 538 DPKTEFEVVDGKLCCTSKPTSQVPLAQRYYTF 568
Db 541 DPTEYEVKVDGEMITCKPFEVALAQRYYFLF 571

RESULT 5
D69729
urease (alpha subunit) ureC - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 05-Oct-2004
C:Accession: D69729
R:Kunzt, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea
C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritzt, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Hatwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Toognoni, A.; Tosaato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:93844377
A:Accession: D69729
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KUN>
A:Cross-references: UNIPROT:P77837; UNIPARC:UPI000060B94; GB:Z99122; GB:AL009126; NID:G
A:Experimental source: strain 168
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
F:3-552/Domain: urease 62K chain homology <U62>

Query Match 66.1%; Score 1981.5; DB 2; Length 569;
Best Local Similarity 62.5%; Pred. No. 5-2e-134;
Matches 357; Conservative 95; Mismatches 114; Indels 5; Gaps 3;

Qy 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVHDYTTYGBELKFGAGKTIREGMQGSN-- 58
Db 1 MKMSREYAEFLGPTGDKIRLGDTDLWIEVEKDFTVYGEEMIFGCGKTIRDCMGQNGRI 60
Qy 59 SPDENTLDLVITNAMIIDYGIYKADIGIKNGKIHGIGKAGNKMDOGVSPhMVVGTEAL 118
Db 61 TGKDGALDITVNVLLDITVGIYKADVGKDGIRVGVKSGNPDIMDGVDPHNVIGAGTE 120
Qy 119 ALAGEGMITAGGIDSHTHFLSPQPPPTALANGVTMTFGGGTGPVDTGNATTITPPGKNL 178
Db 121 VTSGEKLITAGGVDTTHIFICPQQNEVALSSGVTTLLGGTGTPATGSKATTCTSGAWY 180
Qy 179 HRMLRAAEYSVMNVGLFGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ARMLEAAEFPPINVGLFGKNASDKAPLIEQVEAGAIGLKLHEDWGTTPSAINTCEVVD 240
Qy 239 EYDVQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPDVTWAGELNLPSS 298
Db 241 EADIQVAIHTDTNEAGFLNEDLDAIGDRVITHYHIEGAGGHPDITMKLASVANILPSS 300
Qy 299 TPTPTPTINTVAEHLDMLTCHHDKRREDIQLFSQSRIRPGSIAAEDVLHDMGVIAMT 358

Db 301 TPTPTPTINTVAEHLDMLTCHHDKRREDIQLFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360
Qy 359 SSDSQMGRVGEVIIRTWQTADMKKQKORALAEQQKGNNDNVRIRKYVSKYTIINPAITHG 417
Db 361 SSDSQMGRVGEVIIRTWQTADMKKQKORALAEQQKGNNDNVRIRKYVSKYTIINPAITHG 418
Qy 418 VSEYIGSVEEGKIADLVVWNPAGFFGKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREM 477
Db 419 LSHEVGSVEGKLADLVWDPFRFGVKPELILKGLIAYSQMGDPNASIPTPOPVSRRPM 478
Qy 478 FGHGKAKFDTSTITFVSKVAYENGVEKLGLEBRQVLVVKNCRNITKKDKFKNDKTAKITV 537
Db 479 YASYGKANRSTSTITFMSQASIERGVAESLGEKRIKSPVNRIRKSLDKMLNSALPKIEI 538
Qy 538 DPKTEFEVVDGKLCCTSKPTSQVPLAQRYYTF 568
Db 539 DPKTYQVADGEELSCQPDVYVPLGQRYFLF 569

RESULT 6
S75169
urease (EC 3.5.1.5) alpha chain - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1750
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shampo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75169
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KAN>
A:Cross-references: UNIPROT:P73061; UNIPARC:UPI0000137D79; EMBL:D90903; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 64.3%; Score 1927.5; DB 2; Length 569;
Best Local Similarity 62.3%; Pred. No. 3.8e-130;
Matches 354; Conservative 97; Mismatches 114; Indels 3; Gaps 2;

Qy 2 MKKKQEVYNTYGTGDKVRLGDTDLWAEVHDYTTYGBELKFGAGKTIREGMQGSN-SP 60
Db 4 RMDRHSYAHTFGPTGVDKVRADLTFLFEVQDYATYGVDEYKFGGKVRIDGMGQSPSLR 63
Qy 61 DENTLDLVITNAMIIDYGIYKADIGIKNGKIHGIGKAGNKMDOGVSPhMVVGTEAL 120
Db 64 AEGAVDVVITNALLDWMGIVKADVGKNGRIYAGKAGNPHIQDNVS--IIIGPSTBAI 121
Qy 121 AGEGMITAGGIDSHTHFLSPQPPPTALANGVTMTFGGGTGPVDTGNATTITPPGKNLHR 180
Db 122 AGEGMILTAGGIDAHVHFICPQIETALASGVTTLVGGTGTPAAGTKATTCTPGAWNHR 181
Qy 181 MLRAAEYSVMNVGLFGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
Db 182 MLQAADGFPINLGLFGKNGSQPAALAEQIKAGAIKGLKHLHEDWGTTPAIDNCLGVAEDY 241
Qy 241 DVQVCIHTDVTNEAGYVDDTLNANNGRAIHAYHIEGAGGHSPPDVTWAGELNLPSS 300
Db 242 DVQVAIHTDVTNEAGFVETTTAALKNVIHTYHTEGAGGHPDIIKTCGELNLPSS 301
Qy 301 PTPPTPTINTVAEHLDMLTCHHDKRREDIQLFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360

Db 302 PTPPYTINTLLEHDLMLVCHHLHRLNPEDVAFASERIRRETTAAEDILHDLGAFSISS 361
Qy 361 DSQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYSKYTINPALTHGVSE 420
Db 362 DSQAMGRVGEVICRTWTQAHKMKVQRGQLPGETGNNDNLRAKYVAKYTINPALTHGISE 421
Qy 421 YIGSVEEGKIADLVVNPAPFVGKPKIIVKGMVVFSEMGDSNASVPTPQPVYVREMFGH 480
Db 422 EIGSVEVGKLADLCLMSPAFVGKPELVIRKGGIVAVAQMGDANASIPTPQPVHMQPMFAN 481
Qy 481 HGKAKFDTSTTFVSKVAYENGVEKLGRLERQVLPVKNCRNITKKDFKFNDKTAKITVDPK 540
Db 482 YGGAIAATSTTFVSKRAKXDIIGKLGSLKPLVAVRNIRQLTKRDLKLNLYLPHIEVDPE 541
Qy 541 TPEVFDGKLGKTSKPTSOVPLAQRYTF 568
Db 542 TYEVADGELLTCEPASVLPMAQRYFLF 569

RESULT 7
AG2264
urasee alpha chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 05-Oct-2004
C:Accession: AG2264
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; UID:21595285; PMID:11759840
A:Accession: AG2264
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <KUR>
A:Cross-references: UNIPARC:UPI00000CE887; GB:BA000019; PIDN:BA875369.1; PID:g17132803;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3670
C:Superfamily: urease, alpha subunit; urease 62K chain homology

Query Match 63.3%; Score 1897; DB 2; Length 568;
Best Local Similarity 61.4%; Pred. No. 5.8e-128;
Matches 349; Conservative 95; Mismatches 120; Indels 4; Gaps 3;

Qy 2 KMKKQYVNTYGTGDKVRLGDTDLMAEVEHDYTTTVEELKFGAGKTIREGMQSNPD 61
Db 4 RMSRQVAYETGPTVGRIRLADTELFIQVEQDFTTYGDEVKFGGKVIKRGMSQSPIAN 63
Qy 62 -ENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGNKQMDGVSPHVVGVGTAL 120
Db 64 ADGAVDLVTNALILDMWGVKADIGIKNGKIHGKAGNPIYQDHVD--IIIGPGTEAL 121
Qy 121 AGSGMIIITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTITPGKNLHR 180
Db 122 AGSGMIIITAGGIDTHIFICPQQLVAVASGITTMIGGGTGPATGNATTCTPGPMNIR 181
Qy 181 MLRAAEYSNMNVGFLGKGNSSKKQVQVEAGAIGFKLHEDGTTTPSAIDHCLSVAD 240
Db 182 MLQAADAPFNGLGFLGKNASQPGVLEQIFAGAIKGLHEDGTTTPATIDTCLTVADEY 241
Qy 241 DVQVCIHTDVTNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSTT 300
Db 242 DVQVAIHTDVLNAGFVEDTIAAFKNRAIHTYHTEGAGGGHAPDIIKVCQANVLPSTN 301
Qy 301 PTPPYTINTVAEHLMDLMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSS 360
Db 302 PTPPYTINTLDEHDLMLVCHHLDPAAEDVAFASERIRRETTAAEDILHDLGAFSMIAS 361
Qy 361 DSQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYSKYTINPALTHGVSE 420
Db 362 DSQAMGRVGEVIIRPTWTQSHKMKVQSGSLTGDA--EADNLRAKYVAKYTINPALTHGIAQ 420

Qy 421 YIGSVEEGKIADLVVNPAPFVGKPKIIVKGMVVFSEMGDSNASVPTPQPVYVREMFGH 480
Db 421 YVGSVEAGKLADLCLMSPAFVGKPELVIRKGGIAMSQMGDANASIPTPQPVHMRPMPFGS 480
Qy 481 HGKAKFDTSTTFVSKVAYENGVEKLGRLERQVLPVKNCRNITKKDFKFNDKTAKITVDPK 540
Db 481 FAGARNATSTTFVSOALEREIPQOLGRLKSAVAVSGTRQLTKQDKMLNDALPHIEVDSE 540
Qy 541 TPEVFDGKLGKTSKPTSOVPLAQRYTF 568
Db 541 SYEVADGELLTCEPATVLPMAQRYFLF 568

RESULT 8
S47104
urease (EC 3.5.1.5) 62K chain - Bacillus pasteurii
N:Alternate names: ureC protein
C:Species: Bacillus pasteurii
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S47104
R:Moersdorf, G.; Weinmann, P.; Kaltwasser, H.
submitted to the EMBL Data Library, June 1994
A:Description: Nucleotide sequence of three genes on a urease encoding DNA-fragment from Bacillus pasteurii
A:Reference number: S47102
A:Accession: S47104
A:Molecule type: DNA
A:Residues: 1-569 <MOE>
A:Cross-references: UNIPROT:P41020; UNIPARC:UPI0000137D66; EMBL:X78411; NID:g498708; PID:g498708;
C:Genetics:
A:Gene: ureC
A:Start codon: GTG
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterodcamer; hydrolase; metalloprotein; nickel
F:3-552/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted

Query Match 63.0%; Score 1889.5; DB 2; Length 569;
Best Local Similarity 59.1%; Pred. No. 2e-127;
Matches 337; Conservative 112; Mismatches 118; Indels 3; Gaps 2;

Qy 1 MMKKQYVNTYGTGDKVRLGDTDLMAEVEHDYTTTVEELKFGAGKTIREGMQSN-- 58
Db 1 MKINRQYAESYGTGDRVRLADTDL-GEVEKDYIYLGDEVNFGGKVLREGMGNGTY 59
Qy 59 SPDENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGNKQMDGVSPHVVGVGT 118
Db 60 TRTENVDLTLNALILDYTGKADIGKVGKDIYVIGIKGGNPDIMDGVTPNNIVGTATE 119
Qy 119 ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGTNATTITPGKNL 178
Db 120 VIAAEGKIITAGGIDTHVHFNPQDVVALANGITTLFGGGTGAEGSKATTVTGPWN 179
Qy 179 HRMLRAAEYSNMNVGFLGKGNSSKKQVQVEAGAIGFKLHEDGTTTPSAIDHCLSVAD 238
Db 180 EKMLKSTEGLPINVGILGKHGSSIAPIEIQIDACAAGLKIHEDWGATPASIDRSLTVD 239
Qy 239 EYDVQVCIHTDVTNAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSS 298
Db 240 EADVQVAIHTDVLNAGFLEDVTVAINGRVVHSHFVEGAGGGHAPDIMAGHPNVLPS 299
Qy 299 TPTPIPTINTVAEHLMDLMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMT 358
Db 300 TNPTRPTVTNTIDSHLDMLVCHHLKQNPIDVAFADSRIRPETIAAEDILHDLGIISM 359
Qy 359 SSDSQAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYSKYTINPALTHGV 418
Db 360 STDALAMGRAGEVIRPTWTQADKNKKEFGKLPEDGKNDNFRIKRYSKYTINPALAQCM 419
Qy 419 SEYTSVEEGKIADLVVNPAPFVGKPKIIVKGMVVFSEMGDSNASVPTPQPVYVREM 478
Db 420 AHEVGSIEEGKADLVLEWPEKFFGVKADRVIKGGIIAYAQIGDPSASIPTPQPVMMRRY 479
Qy 479 GHGKAKFDTSTTFVSKVAYENGVEKLGRLERQVLPVKNCRNITKKDFKFNDKTAKITVD 538

Db	480	GTVGDLIHDNTITFMSKSIQQGVPAKLGRIGTVKNCRNIGKDKMKWNDVTTDIDIN	539
Qy	539	PKTFEVFDGKLGCTSKPTSOVPLAQRYPF 568	
Db	540	PETYEVKVDGEVLTCEPVKELPMAQRYPFLP 569	
RESULT 9			
H83037			
urease alpha subunit PA4868 [imported] - Pseudomonas aeruginosa (strain PA01)			
C:Species: Pseudomonas aeruginosa			
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 06-Jan-2003			
C:Accession: H83037			
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz			
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,			
.; Lory, S.; Olson, M.V.			
Nature 406, 959-964, 2000			
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho			
A:Reference number: A82950; MUID:20437337; PMID:10984043			
A:Accession: H83037			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-566 <STO>			
A:Cross-references: UNIPARC:UPI00000C5E4A; GB:AE004900; GB:AE004091; NID:g9951134; PIDN:			
A:Experimental source: strain PA01			
C:Genetics:			
A:Gene: ureC; PA4868			
C:Superfamily: urease, alpha subunit; urease 62K chain homology			
Query Match 61.1%; Score 1832; DB 2; Length 566;			
Best Local Similarity 59.2%; Pred. No. 2.6e-123;			
Matches 336; Conservative 88; Mismatches 142; Indels 2; Gaps 1;			
Qy	1	MMKKQEVNTYGTGKGVRLGDTDLWAEVEHDYTYGEBLKFAGKTIREGMGQSNP 60	
Db	1	MKISRQAYADMGFTGVDRLADTLTVEVERDFTYGEVYKFGGKVIKIDMGQSQLG 60	
Qy	61	DENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGKNDMDGVSPHVVGVCTEAL 120	
Db	61	AAQVDVTVTNALILDHMGVAVKADVLGKGRQAIGKAGNPDIQPGV--NIAIGAGTEVI 118	
Qy	121	AGEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTGNATTITPGKNLHR 180	
Db	119	AGEGMIITAGGIDTHLHFICPQIEALMSGVTTMIGGCTGPAAGTNAITCTSGPWHMAR 178	
Qy	181	MLRAAEYSNMVGLGKGNSSKKQLVQVEAGAIGFKLHEDMGTTPSAIDHCLSLVADEY 240	
Db	179	MLQAADAFPMNIGFTGKGNASLPLEEQVLAGAIGLKLHEDMGSTPAAIDNCLVEAERH 238	
Qy	241	DVQVCIHTDVTNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTWAGELNLPSSTT 300	
Db	239	DIQVAIHTDNLNESGEVETTLGAFKRTIHTVHTEGAGGGHAPDIIKACGFANVLPSSTN 298	
Qy	301	PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAITSS 360	
Db	299	PTRPFTNTVIDEHLDMVLMVCHLDPALAEADVAFAESRIRETIAAEDILHDLGAFNMIS 358	
Qy	361	DSQAMGRAGEVITPRTWQADKNKKEFGKLPEDGKNDNFRPRIKYISKYTIINPALTHGVSE 420	
Db	359	DSQAMGRVGEVITRTWQADKNKQRRGLDGDGARDNFRARIYIAKYTIINPAITHGISH 418	
Qy	421	YIGSVEEGKIADLVVWNPAPFGVKPIVIGKGMVFESEMGDSNASVPTPQPVVYREMPGH 480	
Db	419	EYGSVEAGKWADLVLRPAPFGVKPSLILKGAIAASLMGIDNGSITPTPQPVHYRPMFAS 478	
Qy	481	HGKAKFPDTSITFVSKVAYENGKVKGLERQVLVPVNCNITKKDKFNDKTAKITVDPK 540	
Db	479	YAGSRHATSLITFVSKAAFAAGVPPQQLGLRKAIGVSGCRGVQKTDLIHNGLYLTIEVDAQ 538	
Qy	541	TFEVFDGKLGCTSKPTSOVPLAQRYPF 568	
Db	539	NYQVRADGQLWCPEADVLPMAQRYPFLP 566	

RESULT 10			
C75586			
urease, alpha subunit - Deinococcus radiodurans (strain R1)			
C:Species: deinococcus radiodurans			
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004			
C:Accession: C75586			
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; P			
.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma			
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.			
Science 286, 1571-1577, 1999			
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.			
A:Reference number: A75250; MUID:20036896; PMID:10567266			
A:Accession: C75586			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-568 <WHI>			
A:Cross-references: UNIPARC:UPI00000C1687; GB:AE001863; GB:AE001825; NID:g6460670; PIDN:			
A:Experimental source: strain R1			
C:Genetics:			
A:Gene: DRA0318			
A:Map position: 2			
C:Superfamily: urease, alpha subunit; urease 62K chain homology			
Query Match 61.1%; Score 1832; DB 2; Length 568;			
Best Local Similarity 59.0%; Pred. No. 2.7e-123;			
Matches 337; Conservative 100; Mismatches 128; Indels 6; Gaps 3;			
Qy	1	MMKKQEVNTYGTGKGVRLGDTDLWAEVEHDYTYGEBLKFAGKTIREGMGQSN- 59	
Db	1	MKVSROQYADLYGPTGVDRLGDTTELLIEVERDLTTYGEVYKFGGKVIKIDGLQSSAA 60	
Qy	60	--PDENTLDLVTNAMIIDYTGKADIGIKNGKIHGKAGKNDMDGVSPHVVGVGT 117	
Db	61	TRDANVPDLVITNALILDYTWGVTKADVGKNGKRISAIGKAGNPGTQGVTPGLTIAAST 120	
Qy	118	EALAGEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGCTGPDGTGNATTITPGKN 177	
Db	121	EIVAGGLVLTAGVDTHIFIAFQQCTWALESGVTMTTGGTGCTAGTSATTCPTGQWH 180	
Qy	178	LHRMLRAAEYSNMVGLGKGNSSKKQLVQVEAGAIGFKLHEDMGTTPSAIDHCLSLVA 237	
Db	181	IHRMLESAGLPLNFGLLGKGNASTQPPLAEQIRAGALGLKLHEDMGTTPAIHAALSLVA 240	
Qy	238	DEYDVQVCIHTDVTNEAGYVDDTLNMGRAIHAYHIEGAGGHSPPDVTWAGELNLP 297	
Db	241	EDYDVQVAIHTDNLNESGFVEDAIRAFAGRTIHTFTHEGAGGGHAPDIIRVAGLPNVLPS 300	
Qy	298	STPTPTPTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAM 357	
Db	301	STNPTWPTNTTIEHLDMLVMCHLSLSPRIPEDVHFAESRIKRIPTIAAEDVLHDLGVFSM 360	
Qy	358	TSSDSQAMGRAGEVITPRTWQADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHG 417	
Db	361	MSSDSQAMGRVGEVITRTWQAAHKMKVQGRPLAPDGR-ADNFRARYVAKYTIINPAIAG 419	
Qy	418	VSEYIGSVEGKIADLVVWNPAPFGVKPIVIGKGMVFESEMGDSNASVPTPQPVYIREM 477	
Db	420	ISHEVSGSVEGKLADLVLSPPAFPGAKPSLILKGLVVAQMGDANASIPTPQPVYPRPM 479	
Qy	478	FGHHGKAKPDTSITFVSKVAYENGKVKGLERQVLVPVNCNITKKDKFNDKTAKITV 537	
Db	480	FAAYGCGPDATCLHFLVFSQAGLEGGHLPDVG--RYSAVKHTRDTRDIOGKMDQLNAETPDIOV 537	
Qy	538	DPKTFEYFDGKLGCTSKPTSOVPLAQRYPF 568	
Db	538	NPETYEVRVNGELVTCPEVDLPLAQKYPFLP 568	
RESULT 11			
H64075			
urease (EC 3.5.1.5) 62K chain - Haemophilus influenzae (strain Rd KW20)			

Db 492 SARHCRITFLSQAAANGVAERLNLRSIAIVVKGCTRVQKADMVHNSLQPNITVDQTY 541
QY 543 EYFVDGKLTCKPTSOVPLAQRYYTF 568
Db 542 EVRVDGELITSEPADVLPMARQRYFLF 567
RESULT 13
S42607
urease (EC 3.5.1.5) - Rhizobium melliloti
C:Species: Rhizobium melliloti
C>Date: 25-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S42607
R:Mikech, G.; Arnold, W.; Lentzsch, P.; Priefer, U.B.; Puehler, A.
Mol. Gen. Genet. 242, 539-550, 1994
A:Title: A 4.6 kb DNA region of Rhizobium melliloti involved in determining urease and by
ing frames.
A:Reference number: S42601; MUID:94166766; PMID:8121412
A:Accession: S42607
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-570 <NIK>
A:Cross-references: UNIPROT:P42885; UNIPARC:UPI000016FFCB; GB:S69145; NID:G545795; PIDN:
C:Genetics:
A:Gene: ureC
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase; metalloprotein; nickel
F:5-53/Domain: urease 62K chain homology <U62>
F:136,138,219,362/Binding site: nickel 2 (His, His, Lys, Asp) #status predicted
Query Match 60.8%; Score 1822; DB 2; Length 570;
Best Local Similarity 59.8%; Pred. No. 1.4e-122;
Matches 340; Conservative 95; Mismatches 130; Indels 4; Gaps 3;
QY 2 KMKQBYVNTYPTKGDVRLGDTDLWAEVEHDYTTYGEEELFKGAGKTIREGQGNSPD 61
Db 4 RMSRAAYANMFGPTGDKVRLADTELPIEVEKDFTHGEEVFGGKVRDGMGQSQVTR 63
QY 62 E-NTLDLVITNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHMVVGTEAL 120
Db 64 EGGAVDTVITNALILDHWGIVKADIGLKGRIRAAIKAGNPNQPGVT--IIVGPGTEVI 121
QY 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGNTATTITPGKWNLR 180
Db 122 AGEGLVITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGNTATTITPGKWNLR 181
QY 181 MLRAAEYSNMVGLFGKGNSSKKQLVEQVGAIGFKLHEDWGTTTPSAIDHCLSLVADEY 240
Db 182 MIEAADAPFNNLAFACKGNASLPALVEMVGLGATSLKLHEDWGTTTPAIDCCLSVADEY 241
QY 241 DVQVCIHDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGHSDDVITMAGELNLPSTTT 300
Db 242 DVQVMHTDNLNBSGFEVDITAAIKGRTHAYHTEGAGGGHAPDVIIRICQPNVIPSSTN 301
QY 301 PTIPYNTVAEHLDMLTCHLHDKRIRDLQFSQSIRPGSIAAEDVLHDMGVIAMTSS 360
Db 302 PTRPYNTVAEHLDMVCHHLSPTIPEDIAFAESRIRKETAAEDILHIDGAFSISS 361
QY 361 DSQAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSE 420
Db 362 DSQAMGRVGEVIRPTWQTADKNKQRLKEEFGKNDNFVRYIAKYITINPALHAGLSH 421
QY 421 YIGSVEGKIADLVVNNPAPFFGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMFH 480
Db 422 EIGSLEVGKADLVNAPAFFGVKPKIVIKGMVLLGGTIAAAPMGDPNAAIPTPQVHYRPMFGA 481
QY 481 HGKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCR-NITKDFKFNDKTAKITVDP 539
Db 482 YGSRSTNSSVTFVSPASLDAGLRGVAKEGLAVQNTRGIGKASMIHNSLTPHIEVDP 541
QY 540 KTEFVFDGKLTCKPTSOVPLAQRYYTF 568

Db 542 ETYEVRADGELLTCBPATVLPMAQRYPFLF 570
RESULT 14
D43719
urease (EC 3.5.1.5) 62K chain - Proteus mirabilis
N:Alternate names: urease alpha chain
C:Species: Proteus mirabilis
C>Date: 03-Mar-1993 #sequence_revision 02-Dec-1994 #text_change 05-Oct-2004
C:Accession: D43719
R:Jones, B.D.; Mobley, H.L.T.
J. Bacteriol. 171, 6414-6422, 1989
A:Title: Proteus mirabilis urease: nucleotide sequence determination and comparison with
A:Reference number: A43719; MUID:90078080; PMID:2687233
A:Accession: D43719
A:Molecule type: DNA
A:Residues: 1-567 <JON>
A:Cross-references: UNIPROT:P17086; UNIPARC:UPI0000137D74; GB:M31834; NID:G150914; PIDN:
C:Genetics:
A:Gene: ureC
C:Complex: trimer of heterotrimeric subunits consisting of 11K (gamma), 12K (beta), and
C:Function:
A:Description: catalyzes the hydrolysis of one molecule of urea to carbon dioxide and two
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: heterotrimer; hydrolase; metalloprotein; nickel
F:4-550/Domain: urease 62K chain homology <U62>
F:134,136,217,360/Binding site: nickel 2 (His, Lys, Asp) #status predicted
F:217,246,272/Binding site: nickel 1 (Lys, His, His) #status predicted
F:217/Binding site: carbon dioxide (Lys) (covalent) (by urease activase) #status predicted
F:219,320/Active site: His #status predicted
Query Match 60.5%; Score 1813; DB 1; Length 567;
Best Local Similarity 59.5%; Pred. No. 6.1e-122;
Matches 337; Conservative 91; Mismatches 136; Indels 2; Gaps 1;
QY 3 MKKQBYVNTYPTKGDVRLGDTDLWAEVEHDYTTYGEEELFKGAGKTIREGQGNSPDE 62
Db 4 ISRAQVADMFPTGDKRLRLADTELFLIEKDFTYGEEVFGGKVRDGMGQSQVWSA 63
QY 63 NTLDLVITNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDGVSPHMVVGTEALAG 122
Db 64 ECDVLITNAILDLYWGIKADIGIKGRIVGIGKAGNPDVQPNVD--IVIGPGTEVVAG 121
QY 123 EGMIIITAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGNTATTITPGKWNLRML 182
Db 122 EGKIVTAGGIDSHTHFLSPQFPPTALANGVTTFMGGGTGPVDGNTATTITPGKWNLRML 181
QY 183 RAAAEYSNMVGLFGKGNSSKKQLVEQVGAIGFKLHEDWGTTTPSAIDHCLSLVADEYDV 242
Db 182 EAVDELPIVNLFGKGCVSQPEAREQITAGAILGKIHEWDGATPMAIHNCINVADEMDV 241
QY 243 QVCIHDTVNEAGYVDDTLNANMGRAIHAYHIEGAGGHSDDVITMAGELNLPSTTPT 302
Db 242 QVAIHSDTLNEGFEYEEYKAIAGRVHVFTEGAGGGHAPDVIKSGVEPNILPASTNPT 301
QY 303 IPTYNTVAEHLDMLTCHLHDKRIRDLQFSQSIRPGSIAAEDVLHDMGVIAMTSSDS 362
Db 302 MPYNTVNTVAEHLDMVCHHLSPTIPEDIAFAESRIRKETAAEDILHDMGAISSVSSDS 361
QY 363 QAMGRAGEVIRPTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEYI 422
Db 362 QAMGRVGEVIRPTWQCAHKMKLQRLTLAGDSADNDNRIKRYIAKYITINPALAHGIAHTV 421
QY 423 GSVEEGKIADLVVNNPAPFFGVKPKIVIKGMVMVFSEMGDSNASVPTPQVYVREMFHGH 482
Db 422 GSIEKGKADLVLDPAFFGVKPKALIIKGMVMVAPMGDINAAIPTPQVHYRPMYACIG 481
QY 483 KAKPEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKDFKFNDKTAKITVDPKTF 542
Db 482 KAKYQTSMI FMSKAGIEAGVPEKGLSLIORVEGCRHITKASMIHNNVPHIELDPQTY 541
QY 543 EVFVDGKLTCKPTSOVPLAQRYYTF 568

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:57:42 ; Search time 218.186 Seconds
(without alignments)
1836.687 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQYVNTYPTKGDV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2336	77.6	569	Q8GH97_HELBI	Q8gh97 helicobacte
2	2301	76.7	569	Q7X3W5_HELPY	Q7x3w5 helicobacte
3	2298.5	76.6	568	1_URE2_HELHE	P42823 helicobacte
4	2297	76.6	569	1_URE2_HELPY	P69997 helicobacte
5	2297	76.6	569	1_URE2_HELPY	P69997 helicobacte
6	2297	76.6	569	2_Q9SQO5_HELPY	Q9s0q5 helicobacte
7	2292	76.4	567	2_Q64BY3_HELPY	O64ey3 helicobacte
8	2291	76.4	569	2_Q9AFB1_HELPY	Q9afb1 helicobacte
9	2288	76.3	569	2_Q84F75_HELPY	Q84f75 helicobacte
10	2251	75.1	559	2_Q8RNU6_HELPY	Q8rnu6 helicobacte
11	2244	74.8	569	1_URE2_HELFE	Q08716 helicobacte
12	2107	70.3	569	2_Q93R74_HELPHP	Q93rp4 helicobacte
13	2028.5	67.6	569	1_URE1_BACSB	O07397 bacillus sp
14	2016.5	67.2	571	2_Q9K59_BACHD	Q9k59 bacillus ha
15	1981.5	66.1	569	1_URE1_BACSU	P77837 bacillus su
16	1978.5	66.0	569	2_Q5KYM1_GEOKA	O5kym1 geobacillus
17	1960.5	65.4	569	2_Q52305_SYN2P	O52305 synecococc
18	1927.5	64.3	569	1_URE1_SYNY3	P73061 synecocyst
19	1921	64.1	568	2_Q4LNH4_9BURK	O4lnh4 burkholderi
20	1919	64.0	570	2_Q733J6_BACCL	Q733j6 bacillus ce
21	1914	63.8	566	2_Q4ZN06_PSESM	Q4zn06 pseudomonas
22	1914	63.8	566	2_Q87VP0_PSESM	Q87vp0 pseudomonas
23	1907	63.6	572	2_Q8DMV6_SYNEL	O8dmv6 synecococc
24	1906	63.6	565	2_Q5PR23_CAMLA	O5fb23 campylobact
25	1900	63.4	568	2_Q63RL3_BURFS	O63rl3 burkholderi
26	1897	63.3	568	2_Q62HS0_BURMA	Q62hs0 burkholderi
27	1897	63.3	568	2_Q8YQZ0_ANASP	O8yqz0 anabaena sp
28	1896	63.2	566	2_Q4KJ10_PSEF5	O4kj10 pseudomonas
29	1894	63.2	566	2_Q4IXD2_AZOVI	O4ixd2 azotobacter
30	1889.5	63.0	569	1_URE1_BACPA	P41020 bacillus pa
31	1873.5	62.5	569	2_Q9L644_PROWA	Q9l644 prochloroco

32	1873.5	62.5	569	2_Q7V1B6_PROMP	Q7v1b6 prochloroco
33	1865.5	62.2	569	2_Q7U3I3_SYNXP	Q7u3i3 synecococc
34	1854.5	61.8	569	2_Q87402_SYNECO	O87402 synecococc
35	1853.5	61.8	572	2_Q89UG0_BRAJA	O89ug0 bradyrhizob
36	1852.5	61.8	570	2_Q30337_ALCEU	Q30337 alcaligenes
37	1846	61.6	574	2_Q7V3V2_PROMM	Q7v3v2 prochloroco
38	1843.5	61.5	569	2_Q5LSQ2_SILPO	O5lsq2 silicibacte
39	1843	61.5	570	2_Q8RPY2_RHILV	Q8rpy2 rhizobium l
40	1833	61.1	572	2_Q8XXT1_RALSO	Q8xxxt1 ralstonia s
41	1832	61.1	566	2_Q9HUU5_PSEAE	Q9huu5 pseudomonas
42	1832	61.1	568	2_Q9RYJ4_DEIRA	Q9ryj4 deinococcus
43	1832	61.1	572	2_Q4QN09_HAEI8	Q4qn09 haemophilus
44	1831	61.1	567	2_Q5E728_VIBF1	Q5e728 vibrio fisc
45	1831	61.1	837	2_Q949H4_SOYBN	Q949h4 glycine max

ALIGNMENTS

RESULT 1
Q8GH97_HELBI PRELIMINARY; PRT; 569 AA.
AC Q8GH97;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Urease B.
OS Helicobacter bizzozeronii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=56877;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22540228; PubMed=12652903; DOI=10.1080/1042517021000039230;
RA Zhu J., Teng C.H., Chang C.F., Chang C.D., Simpson K.W., Wei C.,
RA McDonough P., McDonough S., Chang Y.F.;
RT "Cloning and characterization of a Helicobacter bizzozeronii urease
gene cluster.";
RL DNA Seq. 13:321-331(2002).
DR EMBL; AF330621; AA015374.1; -; Genomic_DNA.
DR HSP; P14917; 1E9Y.
DR SMR; Q8GH97; 1-569.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR008295; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRINTS; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61922 MW; 3938FFC3E3C9E26 CRC64;

Query Match 77.6%; Score 2326; DB 2; Length 569;
Best Local Similarity 74.1%; Pred. No. 6,5e-151;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

Qy	2	KMKQYVNTYPTKGDVRLGDTDLMAEVSHDYTYGEELKFGAGKITREGMGQNSPD	61
Db	3	KISRKEVSVYPTTKDKVRLGDTDLILEVEHDCDTTYGEEIKFGGKITRDGMAQTNSPS	62
Qy	62	ENTLDLVITNMIIDTYGIYKADIGIKNGKIHGTCAGNKMDCGVSPHVVGVGTEALA	121
Db	63	SHELDLVLTNALIVDYTYGIYKADIGIKNGKIHGTCAGNKMDCGVSNLCLVGPATEALA	122
Qy	122	GEGMIIITAGGIDSTHFLSPQOFTALANGVTTFMFGGCTGPDVDTNATITPGKWLHRM	181

```
Db 123 ABGLIVTAGGIDTHIHFIPTAFASGITMTIMGGTGADGTNATTITPGRNWLKTM 182
Qy 182 LRAAEYSNMVGLGKGNSSKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRASEYAMNLYLCKGNVSEPSLVQLEAGAIQFKIHEDWGSTPAAIYHCLNVADKYD 242
Qy 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHI EGAGGHS PDVITMAGELNLPSSSTP 301
Db 243 VQVAIHTDTLNEAGCVEDTLQIAGRTIHTFTEGAGGHPADVIKMSGEFNILPASTNP 302
Qy 302 TIPYTIINTVAEHLDMMLTCHHLDKIRLEDLQFSQSRIIRPGSIAAEDVLHDMGVIAVTSSD 361
Db 303 TIPFTVNTAEHMDMLVCHLHDKIKEDVQADSRIRPOTIAAEDTLHDMGIFSISSD 362
Qy 362 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPAIHGISEY 422
Qy 422 IGSEEGKIADLVVNNPAPFGVKPVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 481
Db 423 VGSVEVGKADLVLSPPAFGKIPNNMIKGGFIALSQMGDANASITPTQPVYVREMFHGH 482
Qy 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITTKDKPFNDKTAKITVDPKT 541
Db 483 GKAKEDTNIITFVSQVAYDNGIKBELGLQRLVLPVKNCRNITTKDLKFNDVTAHIEVNPET 542
Qy 542 FEVFDGKLGKTSKPTSQVPLAQRYTFF 568
Db 543 YKVKVDGKEVTSKAADKISLAQLYNLF 569

RESULT 2
Q7X3W5_HELPY
ID Q7X3W5_HELPY PRELIMINARY; PRT; 569 AA.
AC Q7X3W5;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Urease B.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC NCBIMOL:U00096.1
RA Duan G.C., Dai L.P.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY295085; AAP51176.1; -; Genomic_DNA.
DR HSSP; P14917; 1E9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphaN.
DR InterPro; IPR011612; Urease_alphaN.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRfams; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61670 MW; 4C96852826C7A784 CRC64;

Query Match 76.7%; Score 2301; DB 2; Length 569;
Best Local Similarity 73.5%; Pred. No. 3.4e-149;
Matches 417; Conservative 75; Mismatches 75; Indels 0; Gaps 0;

Qy 2 KMKKQBYVNTYGTGDKVRLGDTDLIAEVEHDYTYIGEELKFGAGKTIREGMSQNSPD 61
```

```
Db 3 KISRKEYSVMYGPTTGDVRLGDTDLIAEVEHDYTYIGEELKFGAGKTIREGMSQNSPS 62
Qy 62 ENTLDLIVITNAMIIDYTYGIYKADIGIKNGKHGIGKAGNKMDDGVSPHVVGVTEALA 121
Db 63 KEELDIITNALLVDYTYGIYKADIGIKNGKHGIGKAGNKMDDGVSPHVVGVTEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTMTMFGGGTGVVDGTNATTITPGRNWLKTM 181
Db 123 GEGLIIVTAGGIDTHIHFIPTAFASGITMTIMGGTGADGTNATTITPGRNWLKTM 182
Qy 182 LRAAEYSNMVGLGKGNSSKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Db 183 LRAAEYSNMVGLGKGNSSKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 242
Qy 242 VQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHI EGAGGHS PDVITMAGELNLPSSSTP 301
Db 243 VQVAIHTDTLNEAGCVEDTLQIAGRTIHTFTEGAGGHPADVIKMSGEFNILPASTNP 302
Qy 302 TIPYTIINTVAEHLDMMLTCHHLDKIRLEDLQFSQSRIIRPGSIAAEDVLHDMGVIAVTSSD 361
Db 303 TIPFTVNTAEHMDMLVCHLHDKIKEDVQADSRIRPOTIAAEDTLHDMGIFSISSD 362
Qy 362 SQAMGRAGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLPEEKDNDNFRIKRYISKYTINPAIHGISEY 422
Qy 422 IGSEEGKIADLVVNNPAPFGVKPVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 481
Db 423 VGSVEVGKADLVLSPPAFGKIPNNMIKGGFIALSQMGDANASITPTQPVYVREMFHGH 482
Qy 482 GKAKEDTSITFVSKVAYENGVEKGLERQVLVKNCRNITTKDKPFNDKTAKITVDPKT 541
Db 483 GKAKEDTNIITFVSQVAYDNGIKBELGLERQVLVKNCRNITTKDMQFNDVTAHIEVNPET 542
Qy 542 FEVFDGKLGKTSKPTSQVPLAQRYTFF 568
Db 543 YKVKVDGKEVTSKPTKVSQAQLPSIF 569

RESULT 3
URE2_HELHE
ID URE2_HELHE STANDARD; PRT; 568 AA.
AC P42823;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB;
OS Helicobacter heilmannii.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=35817;
RN (1)
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=2;
RX MEDLINE=94222523; PubMed=8168924;
RA Solnick J.V., O'Rourke J., Lee A., Tompkins L.S.;
RT "Molecular analysis of urease genes from a newly identified uncultured
RL Infect. Immun. 62:1631-1638(1994).
CC -I- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -I- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -I- SUBUNIT: Consists of two subunits (alpha and beta).
CC -I- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -I- SIMILARITY: Belongs to the urease family.
CC -I- CAUTION: In Helicobacter the beta subunit is what is known in
CC other bacteria as the alpha subunit.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

use as long as its content is in no way modified and this statement is not removed.

EMBL; L25079; AAA65723.1; -; Genomic_DNA.
HSSP; P14917; LE9Z.
InterPro; IPR006680; Amidohydro_1.
InterPro; IPR005848; Urease_alpha.
InterPro; IPR011612; Urease_alpha_N.
InterPro; IPR008295; Urease_alphalone.
Pfam; PF01979; Amidohydro_1; 1.
Pfam; PF00449; Urease_alpha; 1.
PIRSF; PIRSF001226; Urease_alpha; 1.
PRINTS; PR01752; UREASE.
TIGRFAMs; TIGR01792; urease_alpha; 1.
PROSITE; PS01120; UREASE_1; 1.
PROSITE; PS00145; UREASE_2; 1.
Hydrolase; Metal-binding; Nickel.
ACT_SITE 321 321 Proton donor (By similarity).
METAL 136 136 Nickel 2 (By similarity).
METAL 138 138 Nickel 2 (By similarity).
METAL 219 219 Nickel 1 and 2 (By similarity).
METAL 248 248 Nickel 1 (By similarity).
METAL 274 274 Nickel 1 (By similarity).
METAL 361 361 Nickel 2 (By similarity).
BINDING 221 221 Substrate (By similarity).
SEQUENCE 568 AA; E2C036DF30377917 CRC64;
Query Match 76.6%; Score 2298.5; DB 1; Length 568;
Best Local Similarity 73.4%; Pred. No. 5e-149;
Matches 418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;
QY 2 KMKKQEVNVTYGTGDKVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEVSMYGTGDKVRLGDTDLWAEVHDCTTYGEEKFGAGKTIREGMQSNPS 62
QY 62 ENTLDIVTNAMIDVTGYIKADIGIKNGKIHGKAGNKMDDGVSPVMVVGTEALA 121
DB 63 SHELDIVTNALIVDTGYIKADIGIKNGKIHGKAGNKLDDGVNRLCVGPATEALA 122
QY 122 GGMITAGIDSHTHFLSPQFPPTALANGVTVMFGGTPGVDTNATITPKWNLHRM 181
DB 123 AEGLIVTAGIDITHFISFQPIPTAFASGITTMIGGTPGADGTNATITPGRWNLKEM 182
QY 182 LRAAEYSNNVGLFGKNSSSKQLVEQVAGAGKFLHEDWGTTPSAIDHCLSVADEYD 241
DB 183 LRASEYAMNLGLGKNSVSPFALDQLEAGAIGFKIHEDWGSTPSAINHALNIADKYD 242
QY 242 VQVCIHDTVNBAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTWAGELNLPSTTP 301
DB 243 VQVAIHDTLNBAGCVEDTLEAIGRTIHTFTGEGAGGGHAPDVIKWAGEFNLPASTNP 302
QY 302 TIPTTINTVAEHLMDLMTCHLDKRIREDLQFSQRSIRPFSIAAEVDLHDMGVIAMTSSD 361
DB 303 TIPTTKNTEAEHMDMLM-CHLDKNIKEDVEFADSRIRPQTIAAEKLDHDMGIFSISSD 361
QY 362 SQAMGRAGEVIRPTWQADKNKPEFKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 421
DB 362 SQAMGRAGEVIRPTWQADKNKPEFKLPEDGKNDNFRIKRYISKYITNPALTHGISEY 421
QY 422 IGSVEGKTDADLVVWNPAPFGVKPKTVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
DB 422 VGSVEGKTDADLVVWNPAPFGVKPKTVIKGMVVFSEMGDSNASVPTPQPVYVREMGHH 481
QY 482 GKAKFTDTSITFSKVAAYENGKELGLERQVLVVKNCRNITKKDKFKFNDKTAITVDPKT 541
DB 482 GKAKFTDTSITFSKVAAYENGKELGLERQVLVVKNCRNITKKDKFNDVTAHIEVNPET 541
QY 542 FEYFVDGKLCSTKPTSQVPLAQRYTFF 568
DB 542 YKVKVDGNEVTSAAADKLSLAQYLNLF 568

RESULT 4

URE2_HELPJ STANDARD; PRT; 569 AA.
ID URE2_HELPJ
AC P69997; P14917;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureb; Synonyms=hpuB; OrderedLocusNames=JHP0067;
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_Taxid=85963;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=99120557; PubMed=9923682; DOI=10.1038/16495;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Heterohexamer of 3 alpha and 3 beta subunits. Four
heterohexamers assemble to form a 16 nm spherical complex (By
similarity).
CC -!- FM: Lys-219 is carbamylated. The carbamoyl group provides the
ligands for the two nickel ions (By similarity).
CC -!- MISCELLANEOUS: Oligomerization may protect the enzyme against
denaturation in an acidic environment (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
other bacteria as the alpha subunit.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
CC EMBL; AE001446; AAD05651.1; -; Genomic_DNA.
DR EIR; B38537; URKCBP
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR InterPro; IPR008295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alpha; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase; Metal-binding; Nickel.
FT ACT_SITE 322 322 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 and 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61684 MW; 4C8A6BC6C8295584 CRC64;
Query Match 76.6%; Score 2297; DB 1; Length 569;
Best Local Similarity 73.4%; Pred. No. 6.3e-149;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
QY 2 KMKKQEVNVTYGTGDKVRLGDTDLWAEVHDYTYGEEKFGAGKTIREGMQSNPSD 61
DB 3 KISRKEVSMYGTGDKVRLGDTDLWAEVHDCTTYGEEKFGAGKTIREGMQSNPS 62

the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; M50398; AAA25021.1; -; Genomic DNA.
EMBL; X17079; CAA34933.1; -; Genomic DNA.
EMBL; A08818; CAA00811.1; -; Unassigned DNA.
EMBL; A07398; CAA00663.1; -; Unassigned DNA.
EMBL; A8000529; AAD07143.1; -; Genomic DNA.
EMBL; M84338; -; NOT ANNOTATED_CDS; Genomic DNA.

PIR; B38537; URKCBP.

PDB; 1E9Y; X-ray; B=1-569.

PDB; 1E9Z; X-ray; B=1-569.

TIGR; HP0072; -.

InterPro; IPR006680; Amidohydro 1.

InterPro; IPR005848; Urease alpha.

InterPro; IPR011612; Urease-alpha.N.

InterPro; IPR008295; Urease-alpha1one.

Pfam; PF01979; Amidohydro 1; 1.

Pfam; PF00449; Urease alpha; 1.

PIRSF; PIRSF001226; Urease_alpha; 1.

PRINTS; PR01752; UREASE.

TIGRfam; TIGR01792; urease_alpha; 1.

PROSITE; PS01120; UREASE_1; 1.

PROSITE; PS00145; UREASE_2; 1.

3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;

Metal-binding; Nickel.

ACT_SITE 322 322 Proton donor (Probable).

METAL 136 136 Nickel 2 (By similarity).

METAL 138 138 Nickel 2 (By similarity).

METAL 219 219 Nickel 1 and 2 (By similarity).

METAL 248 248 Nickel 1 (By similarity).

METAL 274 274 Nickel 1 (By similarity).

METAL 362 362 Nickel 2 (By similarity).

BINDING 221 221 Substrate (Probable).

CONFLICT 10 10 V -> A (in Ref. 2).

CONFLICT 19 19 D -> A (in Ref. 2).

CONFLICT 22 22 R -> C (in Ref. 4).

CONFLICT 27 27 D -> C (in Ref. 4).

CONFLICT 104 104 M -> T (in Ref. 2).

CONFLICT 181 181 W -> F (in Ref. 2).

CONFLICT 193 193 L -> F (in Ref. 2).

CONFLICT 218 218 F -> Y (in Ref. 2).

CONFLICT 273 273 F -> L (in Ref. 2).

CONFLICT 540 540 P -> S (in Ref. 2).

CONFLICT 554 554 SPKANKVSLAQLFSIF -> LNQPIK (in Ref. 2).

STRAND 3 5

HELEX 6 13

TURN 17 18

STRAND 20 22

TURN 24 25

STRAND 29 31

STRAND 35 35

TURN 39 40

TURN 47 48

STRAND 53 53

TURN 54 56

STRAND 57 57

TURN 63 64

STRAND 68 77

TURN 78 79

STRAND 80 89

TURN 90 91

STRAND 92 97

STRAND 100 100

TURN 102 104

TURN 106 107

HELEX 110 112

STRAND 113 113

TURN 116 117

STRAND 119 122

TURN 124 125

STRAND 127 130

FT	STRAND	132	138
FT	TURN	142	143
FT	HELEX	144	151
FT	TURN	152	152
FT	STRAND	153	159
FT	HELEX	165	169
FT	HELEX	175	186
FT	TURN	187	188
FT	STRAND	192	198
FT	HELEX	204	212
FT	TURN	213	214
FT	STRAND	217	221
FT	HELEX	222	224
FT	HELEX	228	240
FT	TURN	241	242
FT	STRAND	244	248
FT	TURN	251	252
FT	HELEX	258	265
FT	TURN	266	267
FT	STRAND	270	272
FT	TURN	275	276

Query Match 76.6%; Score 2297; DB 1; Length 569;

Best Local Similarity 73.4%; Pred. No. 6.3e-149; Mismatches 75; Indels 0; Gaps 0;

Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;

Qy 2 KMKQEVNTYKPTKGDVRLGDTDLMAEVEHDYTYGEEKLFGAGKTIREGMGQSNPSD 61

Db 3 KISRKEYVSMYPTTGDKVRLGDTDLIAEVEHDYTYGEEKLFGAGKTIREGMGQSNPS 62

Qy 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGKAGKQMDQGVSPHVVGVGTEALA 121

Db 63 KEELDLITNALIVDYTYGIYKADIGIKNGKIHGKAGKQMDQGVSPHVVGVGTEALA 122

Qy 122 GEGMIITAGGIDSHTHFLSPQFPFALANGVYTMFGGCTGPDVDTNATITTPCKNHLHM 191

Db 123 GEGLVITAGGIDTHIFHSPOQIPTAFASGVYTMFGGCTGPDVDTNATITTPCKNHLHM 192

Qy 182 LRAAEYSNMVGLKGNSSKKQIIVEQVEAGAIKFLHEDWGTTPSAIDHCLSLVADEYD 241

Db 183 LRAAEYSNMVGLKGNSSKKQIIVEQVEAGAIKFLHEDWGTTPSAIDHCLSLVADEYD 242

Qy 242 VQVCIHTDVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVITMAGELNLPSTTP 301

Db 243 VQVAHTDLNEAGCVEDTMAIAGRTMHTFTEGAGGSHAPDIIKVAGEHNILPASTNP 302

Qy 302 TTPVTINTVAHDLMLTCHHLDKRREDLFQSSQSRIRPGSIAADSVLHDMGVIAWTSSD 361

Db 303 TTPVTINTVAHDLMLTCHHLDKRREDLFQSSQSRIRPGSIAADSVLHDMGVIAWTSSD 362

Qy 362 SQAMGRAGEVITPTWQTADKNKBSFGKLPEDGKONDNFRIKRYISKYITINPALTGVSEY 421

Db 363 SQAMGRAGEVITPTWQTADKNKBSFGKLPEDGKONDNFRIKRYISKYITINPALTGVSEY 422

Qy 422 IGSVEGKIADLVVWNPAPFFGVKPKIVIKGGWVVFSEMGDSNASVPTPQVYIREMFGHH 481

Db 423 VGSVEGKVADLVVWNPAPFFGVKPKIVIKGGWVVFSEMGDSNASVPTPQVYIREMFAHH 482

Qy 482 GKAKFDTSITFVSKVAVENGVEKLGRLQVLPVKNCRNITKQPKFNDKTKAKITVDPKT 541

Db 483 GKAKYDANITFVSKVAVENGVEKLGRLQVLPVKNCRNITKQPKFNDKTKAKITVDPKT 542

Qy 542 FEVFDGKLCITSKPTSQVPLAQRVTF 568

Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569

RESULT 6

Q9S0Q5_HELPY

ID Q9S0Q5_HELPY PRELIMINARY; PRT; 569 AA.

AC Q9S0Q5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)


```
Db 362 QAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEYV 421
Qy 423 GSVEEKGADLVVNPAFFGVKPKI VIKGGMVVVFSEMGDSNASVPTQPVPYYREMFGHHG 482
Db 422 GSVEGKADLVVNSPAFFGVKPNMIIKGFIALSQMGDANASIPTPQVPYYREMFAHGG 481
Qy 483 KAKFDTSIITFVSKVAYENGVEKELGLERQVLPVKNCNRTIKDKFKNDKTAKITVDPKTF 542
Db 482 KAKYDANITFVSQAAYDKGIEKELGLERQVLPVKNCNRTIKDKMQFNDTTAHIEVNPETY 541
Qy 543 EYFVKGKCTSKTSQVPLAQRVTF 568
Db 542 HNFVKGKTSKPAKVSLAQLFSIF 567

RESULT 8
Q9AFBI_HELPY
ID Q9AFBI_HELPY PRELIMINARY; PRT; 569 AA.
AC Q9AFBI;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Urease B.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Gerold A., Simpson K., Chang Y.-F.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF352376; AAK32714.1; -; Genomic_DNA.
DR HSSP; P14917; 1E9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease alpha.
DR InterPro; IPR008295; Urease alpha.
DR InterPro; IPR011612; Urease alpha_N.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 569 AA; 61713 MW; C243ABC38C25D187 CRC64;

Query Match 76.4%; Score 2291; DB 2; Length 569;
Best Local Similarity 73.2%; Pred. No. 1.6e-148;
Matches 415; Conservative 76; Mismatches 76; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPSD 61
Db 3 KISRKEYVSMGPTTGDKVRLGDTDLWAEVEHDYTYGEEELKFGGKTLREGMSQSNPS 62
Qy 62 ENTLDLIVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEALA 121
Db 63 KEELDLIITNALIVDVTGIYKADIGIKNGKIHGIGKAGNKMDQGVKNVNSVGPATEALA 122

Qy 122 GEGMIITAGGIDSHTHFLSPQPPPTALANGVTTMFGGGTGFVDGTNATTTTPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPPQIPTAFASGVTMMIGGGTGPADGTNATTTTPGRNLKWM 182
Qy 182 LRAAEYSNNVGLFGKNSSSKQVVEAGIAGFKLHEDWGTTPSAIDHCLTSVADEYD 241
Db 183 LRAAEYSNNLSPLAKGNASNDASLADQIEAGAIGFKIHEDWGTTPSAINHDLVDADKYD 242
Qy 242 VQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGHSVDVITMAGELNITLPSSTTP 301
Db 243 VQVAIHDTDLNEAGCVEDTMAAIAGRTMHTFHTGAGGGHAPDIIKVAGEHNILPASTNP 302
Qy 302 TIPTYTINTVAEHLDMLTCHHLDKRIREDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSSD 361
Db 303 TIPTVTATEAHEMDMLMVCHHLDKSIKEDVQFADSRIRPQTIAAEDTILHDMGIFSISSD 362
Qy 362 SQAMGRAGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEY 421
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEEKGDNDNFRIKRYLSKYTINPAIAHGISEY 422
Qy 422 IGSVEEKGADLVVNPAFFGVKPKI VIKGGMVVVFSEMGDSNASVPTQPVPYYREMFGHH 481
Db 423 VGSVEGKADLVVNSPAFFGVKPNMIIKGFIALSQMGDANASIPTPQVPYYREMFAHGG 482
Qy 482 GKAKFDTSIITFVSKVAYENGVEKELGLERQVLPVKNCNRTIKDKFKNDKTAKITVDPKTF 541
Db 483 GKAKYDANITFVSQAAYDKGIEKELGLERQVLPVKNCNRTIKDKMQFNDTTAHIEVNPETY 542
Qy 542 FEVFDGKCTSKTSQVPLAQRVTF 568
Db 543 YHNFVKGKTSKPAKVSLAQLFSIF 569

RESULT 9
Q84F75_HELPY
ID Q84F75_HELPY PRELIMINARY; PRT; 569 AA.
AC Q84F75;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Urease beta (EC 3.5.1.5).
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Karimi Arzenani M., Mohammadi M.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY227442; AAO334403.1; -; Genomic_DNA.
DR HSSP; P14917; 1E9Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alpha.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Hydrolase.
SQ SEQUENCE 569 AA; 61697 MW; D35279ED45EF712F CRC64;

Query Match 76.3%; Score 2288; DB 2; Length 569;
Best Local Similarity 73.2%; Pred. No. 2.6e-148;
Matches 415; Conservative 75; Mismatches 77; Indels 0; Gaps 0;

Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEELKFGAGKTIREGMQSNPSD 61
Db 3 KISRKEYVSMGPTTGDKVRLGDTDLWAEVEHDYTYGEEELKFGGKTLREGMSQSNPS 62
Qy 62 ENTLDLIVTNAMIIDYTYGIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEALA 121
Db 63 KEELDLIITNALIVDVTGIYKADIGIKNGKIHGIGKAGNKMDQGVKNVNSVGPATEALA 122
```

QY 122 GEGMIITAGGIDSHTHFLSPQOPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 GEGLIIVTAGGIDTHIHFIISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
QY 182 LRAAEYSNNVGLGKGNSSKQLVEQVEAGAIGFKLHEDGTTTSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIGFKIHEDGTTTSAINHALDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTIMAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVKDTMAALAGRTMTFTHTEGAGGGHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYINTVAEHLDMMLTCHLIDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAMTSSD 361
Db 303 TTPFTVNPAAEHMDMLVCHLIDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVITRTWTQADKNKKEFKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWTQADKNKKEFKLKEEGKNDNFRIKRYLSKYITNPALAHGISEY 422
QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHGH 481
Db 423 VGSVEVGKVDLVWSPAPFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVREMFH 482
QY 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKDFKFNDKTAKITVDPKPT 541
Db 483 GKAKYDANITFVSOAAYDKGIKEELGLERQVLPVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPSQVLAQRYTF 568
Db 543 YRVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 10
QBRNU6_HELPY
ID QBRNU6_HELPY PRELIMINARY; PRT; 559 AA.
AC QBRNU6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease B subunit.
GN Name=ureB;
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=20123532; PubMed=10660136;
RA Opazo P., Muller I., Rolian A., Valenzuela P., Yudelevich A.,
Garcia-de la Guardia R., Urra S., Venegas A.;
RT "Serological response to Helicobacter pylori recombinant antigens in
Chilean infected patients with duodenal ulcer, non-ulcer dyspepsia and
gastric cancer.";
RL APMIS 107:1069-1078 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH-CTX1;
RX MEDLINE=2121498; PubMed=12125208;
RA Muller I., Medina-Selby A., Palacios J.L., Martinez P., Opazo P.,
Bruce E., Mancilla M., Valenzuela P., Yudelevich A., Venegas A.;
RT "Cloning and comparison of ten gene sequences of a Chilean H. pylori
strain with other H. pylori strains revealed higher variability for
VacA and CagA virulence factors.";
RL Biol. Res. 35:67-84 (2002).
DR EMBL; AF479026; AAL86896.1; -; Genomic_DNA.
DR HSSP; P14917; 1E9Y.
DR SMR; Q8RNU6; 1-553.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.

DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphalone.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1_1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRF; PRF001226; Urease_alpha; 1.
DR PRF; PRF01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
SQ SEQUENCE 559 AA; 60607 MW; 049CA0F87E2CABID CRC64;
Query Match 75.1%; Score 2251; DB 2; Length 559;
Best Local Similarity 73.9%; Pred. No. 8.8e-146;
Matches 407; Conservative 72; Mismatches 72; Indels 0; Gaps 0;
QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHEHYTTTGBELKFGAGKTIRMGMSQNSPD 61
Db 3 KISRKEYASMYGPTTGDKVRLGDTDLJAEVHEHYTTTGBELKFGGKTIRMGMSQNSPS 62
QY 62 ENTLDLVIITNAMIIDYTIKADIGIKNGKHGIGKAGNKDMQDGVSPHVVGVCTEALA 121
Db 63 KEELDLIIITNALIVDYTIKADIGIKNGKHGIGKAGNKDMQDGVKNLSVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOPPTALANGVTTMFGGGTGPVDTGNATTITPGKWNLRHM 181
Db 123 GEGLIIVTAGGIDTHIHFIISPOQIPTAFASGVTTMIGGGTGPADGTNATTITPGRRNLKWM 182
QY 182 LRAAEYSNNVGLGKGNSSKQLVEQVEAGAIGFKLHEDGTTTSAIDHCLSVADYD 241
Db 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIGFKIHEDGTTTSAINHALDVADKYD 242
QY 242 VQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTIMAGELNLPSTTP 301
Db 243 VQVAIHDTLNEAGCVKDTMAALAGRTMTFTHTEGAGGGHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYINTVAEHLDMMLTCHLIDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVAMTSSD 361
Db 303 TTPFTVNPAAEHMDMLVCHLIDKSIKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
QY 362 SQAMGRAGEVITRTWTQADKNKKEFKLPEDGKNDNFRIKRYISKYITNPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWTQADKNKKEFKLKEEGKNDNFRIKRYLSKYITNPALAHGISEY 422
QY 422 IGSVEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPOPVYVREMFHGH 481
Db 423 VGSVEVGKVDLVWSPAPFGVKPNMIIKGGFIALSQMGDANASITPTPOPVYVREMFH 482
QY 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKDFKFNDKTAKITVDPKPT 541
Db 483 GKAKYDANITFVSOAAYDKGIKEELGLERQVLPVKNCRNITKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCST 552
Db 543 YRVFVDGKEVT 553

RESULT 11
URE2_HELPY
ID URE2_HELPY STANDARD; PRT; 569 AA.
AC Q087I6; Q083B3;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease beta subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureB;
OS Helicobacter felis.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=214;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

RX MEDLINE=94018627; PubMed=8412683;
RA Ferrero R.L., Labigne A.;
RT "Cloning, expression and sequencing of Helicobacter felis urease
RT genes.";
RL Mol. Microbiol. 9:323-333(1993).
RN [2]
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=ATCC 49179;
RX PubMed=1452359;
RA Turbett G.R., Hoef P.B., Horne R., Mee B.J.;
RT "Purification and characterization of the urease enzymes of
RT Helicobacter species from humans and animals.";
RL Infect. Immun. 60:5259-5266(1992).
CC -!- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
CC -!- COFACTOR: Binds 2 nickel ions per subunit (By similarity).
CC -!- SUBUNIT: Consists of two subunits (alpha and beta).
CC -!- PTM: Lys-219 is carbamylated. The carbamoyl group provides the
CC ligands for the two nickel ions (By similarity).
CC -!- SIMILARITY: Belongs to the urease family.
CC -!- CAUTION: In Helicobacter the beta subunit is what is known in
CC other bacteria as the alpha subunit.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X69080; CA448826.1; -; Genomic_DNA.
DR PIR; C49215; C49215.
DR PIR; S35291; S35291.
DR HSP; P14917; 1E9Z.
DR InterPro; IPRO06680; Amidohydro 1.
DR InterPro; IPRO05848; Urease_alpha.
DR InterPro; IPRO11612; Urease_alpha_N.
DR InterPro; IPRO08295; Urease_alphalone.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Direct protein sequencing; Hydrolase; Metal-binding; Nickel.
FT ACT_SITE 322 322 Proton donor (By similarity).
FT METAL 136 136 Nickel 2 (By similarity).
FT METAL 138 138 Nickel 2 (By similarity).
FT METAL 219 219 Nickel 1 and 2 (By similarity).
FT METAL 248 248 Nickel 1 (By similarity).
FT METAL 274 274 Nickel 1 (By similarity).
FT METAL 362 362 Nickel 2 (By similarity).
FT BINDING 221 221 Substrate (By similarity).
SQ SEQUENCE 569 AA; 61703 MW; 66DBB9952CC7CEA3 CRC64;

Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 2.7e-145;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;

QY 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVHEHDYTYTGEELKFTGAGKTIREGNGQSNPSD 61
DB 3 KISRKEYSVSYGPTTGDVRLGDTDLLEVEHDCTTYGEEIKFTGGKTIRDGMSQTSNPS 62
QY 62 ENTLDLVITNAMITDVTGYIKADIGIKNGKIHGIGKAGNKMDDGVSPHVVGVGTEALA 121
DB 63 SYELDLVLTNALIVDVTGYIKADIGIKDGKGIAGIKAGNKMDDGVNLLNLCVGPATEALA 122
QY 122 GEGMIITAGGIDSHTHFLSPQOPTALANGVTTFMFGGTGTPVDTNATTTTCGNWLHRM 181
DB 123 AEGLIIVTAGGIDTHIHFIISQOIPPTAFASGVTTFMIGGTPADGTNATTTTCGRANLKSM 182
QY 182 LRAAEYSNMVGLFGKGNSSKQQLVEQVSAIGAIFKLHEDWGTTPSAIDHCLSVADEYD 241
DB 183 LRAAEYAMNGLFLAKGNVSYSFLSRDQIEAGAIGFKIHEDWGSTPAIHHCLNVADEYD 242

```

```

QY 242 VQVCIHDTTNEAGYVDDTLNMGRAIHAVHIEGAGGSHSPDVTMAGELNILPSSTTP 301
DB 243 VQVAIHTDTLNEAGCVEDTLEAIGARTIHTFHTGAGGSHAPDVVKMAGEFNILPASTNP 302
QY 302 TIPVTINTVAEHLDMATCHHLDKRIEDLOFSQSRIRPGSIAAEDVLHDMGVIAWTSSD 361
DB 303 TIPFTKNTAEAHMDMLMVCCHLDKSIKEDVOFADSRIRPQTIAAEDQLHDMGIFSTSSD 362
QY 362 SQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKDNDFRIKRIYISKYITINPALTHTGVSEY 421
DB 363 SQAMGRGEVITRTWQTADKNKKEFGRLKEEGKDNDFRIKRIYISKYITINPGIAHGISDY 422
QY 422 IGSVEGKIADLVVMPAPFGVKPKIVIKGMVSEMGDSNASVPTQPQVYYREMFHGH 481
DB 423 VGSVEVGKYADLVLSPPAFPGIKPMI IKGGFIALSQMGDANASITPTQPQVYYREMFHGH 482
QY 482 GKAKFDTSITFVSKVAYENGVEKLGRLERQVLVKNCRNITKKDKPKNDKTKITVDPKT 541
DB 483 GKKNFDTNITFVSOAYKAGIKBEILGLDRAAPPVKNCRNITKKDKPKNDKTKITVDPKT 542
QY 542 FEVFDGKLTCKTSPSOVPLAQRTPF 568
DB 543 YKVKVDGKEVTSKAADELSLAQLNLF 569

RESULT 12
Q93PJ4_HELHP PRELIMINARY; PRT; 569 AA.
AC Q93PJ4; Q7BYA2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Urease beta subunit UreB (EC 3.5.1.5).
GN Name=ureb; OrderedLocusNames=RH0408;
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP NUCLEOTIDE SEQUENCE..
RX MEDLINE=21391860; PubMed=11500473;
RX DOI=10.1128/JAI.69.9.5914-5920.2001;
RA Beckwith C.S., McGee D.J., Mobley H.L.T., Riley L.K.;
RT "Cloning, expression, and catalytic activity of Helicobacter hepaticus
RL urease.";
RL Infect. Immun. 69:5914-5920(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954; DOI=10.1073/pnas.1332093100;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droge M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AF326556; AA659199.1; -; Genomic DNA.
DR EMBL; AE017145; AAP7005.1; -; Genomic_DNA.
DR HSP; P14917; 1E9Y.
DR HSP; P18314; 1A5N.
DR SMR; Q93PJ4; 1-569.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPRO0680; Amidohydro 1.
DR InterPro; IPRO05848; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro 1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.

```



```
QY 482 GKAKEDTSITFVSKVAYENGVEKLGLEROVLVYKNCNRTTKKDFKFNKDKTAKITVDPKT 541
Db 483 GKAKYQTSITFVSKVAYENGVEKLGLEQGLKKVKVPVHGIRKULTKKDLILNDKTPKIDVDPQT 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRYYTF 568
Db 543 YEVKVDGLVTCPEAIVPMAQRYYFLF 569

RESULT 14
ID Q9KG59_BACHD PRELIMINARY; PRT; 571 AA.
AC Q9KG59;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5).
GN Name=ureC; OrderedLocusNames=BH0254;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; BA000004; BAB03973.1; -; Genomic_DNA.
DR PIR; F83681; F83681.
DR HSSP; P41020; 1UBP.
DR SMR; Q9KG59; 1-571.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016151; F:nickel ion binding; IEA.
DR GO; GO:0009039; F:urease activity; IEA.
DR GO; GO:0019627; P:urea metabolism; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR005848; Urease_alpha.
DR InterPro; IPR008295; Urease_alphalone.
DR InterPro; IPR011612; Urease_alpha_N.
DR Pfam; PF01979; Amidohydro_1; 1.
DR Pfam; PF00449; Urease_alpha; 1.
DR PIRSF; PIRSF001226; Urease_alpha; 1.
DR PRINTS; PR01752; UREASE.
DR TIGRFAMs; TIGR01792; urease_alph; 1.
DR PROSITE; PS01120; UREASE_1; 1.
DR PROSITE; PS00145; UREASE_2; 1.
KW Complete proteome; Hydrolase.
SQ
Query Match 67.2%; Score 2016.5; DB 2; Length 571;
Best Local Similarity 64.4%; Pred. No. 1.1e-129;
Matches 368; Conservative 88; Mismatches 112; Indels 3; Gaps 2;
QY 1 MKNKKQEVNTYPTGKGVRLGDTDLMAEVEHDYTYGELKFGAGKTIREGMGQS--N 58
Db 1 MKLTRAQHASLYGPTVGDKVRLADTOLLLEIEKDYTYVYDECKFGGKVLKRDGMGQSAVY 60
QY 59 SPDENTLDIVTNAMLIIDYTVIKADIGIKNGKHIGIKAGNKMDDGVSPHVMVGVGTE 118
Db 61 TRDEGVLDIIITNATIIDYTVIKADIGIKDGHIVGIGKGNPDIMDGVESHMIVGASTE 120
QY 119 ALAGEGMIITAGGIDSHTHFLSPQFPPTALANGVTWTFMGGCTGPDVGTNATITTPGKNWL 178
Db 121 ATAGEGLIVTAGIDAHIIHFISFQQIDVAIASIITMLGGTGTPATGATCTTPGKNNI 180
QY 179 HRMLRAAEBSYMNVLFGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 238
Db 181 ERMLEAADAPPVNLGLFGKGNASTPAPLRQIEAGAILGLKLEDWGTTPPAIRTCLSVAD 240
```

```
QY 239 EYDVQVCIHTDVTNAGVYDDTLNAMNGRAIHAYHIEGAGGHSPOVITMAGELNLPSS 298
Db 241 RMDVQVALHTDLTNEAGFVEDTIIKAGDRVHTYHTEGAGGGAHPIMKVAGLPNVLPS 300
QY 299 TTPTIPYINTVAEHLDMLTCHHLDKRIRDLQFSQSRIIPGSAEAEDVLHDMGVIAMT 358
Db 301 TNTPRPTVNTIDHLDMLAVCHHLDPNVDPVAFADSKIRPETIAAEDILQDLGVISMI 360
QY 359 SSDSQAMGRAGEVTPRTWQTADKNKKEGKLPED-GKNDNPRIRIKYISKYINPALTHG 417
Db 361 SSDSQAMGRVGEVIRTWQTADKNKQKORGALEADQKGNNDNVRIRKYSKYINPAITHG 420
QY 418 VSEYIGSVEECKIADLVVNPAFFGVPKPIVIKGGMVVFSEMGDSNASVPTPPQVYREM 477
Db 421 IDDVYGSVEVEGKLADLVLDPRFFGVKPELILKGLIAYSQMGDPNASIPTQPVSFRPM 480
QY 478 FGHGKAKFDTSITFVSKVAYENGVEKLGLEQVLVYKNCNRTTKKDFKFNKDKTAKITV 537
Db 481 FGSFGRAGTTSITFSLKAAMDVLGVHEALGLQKIAHVKNCSISKRSKMYNDATPNIEI 540
QY 538 DPKTFEVFDGKLCSTKPTSQVPLAQRYYTF 568
Db 541 DPETVEVKVDGEMITCKPFEVALAQRYYFLF 571

RESULT 15
URE1_BACSU
ID_URE1_BACSU STANDARD; PRT; 569 AA.
AC P77837;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase).
GN Name=ureC; OrderedLocusNames=BSU36640;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=168;
RA Glaser P., Cruz H., Danchin A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holzapfel S., Hosono S., Hullo M.-F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Maue C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone P., Seguchi J., Sekowska A., Seror S.J., Serror P.,
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Tognoni K., Takeuchi M., Takakoshi A., Tanaka T., Terpstra P.,
RA Tognoni K., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 07:55:02 ; Search time 185.28 Seconds
(without alignments)
1346.976 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQYVNTYPTKGDV.....KLCTSKPTSQVPLAQRYYTF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2999	100.0	568	5	ADJ58239 Urease su
2	2973	99.1	568	5	Adj58251 UreaseXY
3	2960	98.7	568	5	Adj58242 UreaseXY
4	2953	98.5	568	5	Adj58248 UreaseXY
5	2588	86.3	496	5	ADJ58245 UreaseXY
6	2326	77.6	569	8	Adq37851 H. bizoz
7	2297	76.6	569	2	AAR4580 Part of p
8	2297	76.6	569	2	Aaw07194 H. pylori
9	2297	76.6	569	2	Aaw98511 H. pylori
10	2297	76.6	569	6	ABU30687 Protein e
11	2297	76.6	569	8	ADU05536 H. pylori
12	2282	76.1	568	8	ADS09177 H. pylori
13	2281	76.1	569	2	AAR67378 H. pylori
14	2281	76.1	570	4	AEBS5116 Helicobac
15	2244	74.8	569	2	AAR74337 Helicobac
16	2244	74.8	569	2	Aaw06730 H. felis
17	2238	74.6	749	2	AAR80599 Helicobac
18	2237	74.6	806	2	AAR67371 Urease A
19	2233	74.5	559	2	AAR13550 B subunit
20	2225	74.2	559	8	Adm28644 Helicobac
21	2013.5	67.1	568	2	AAR36387 Urease ga
22	1922	64.1	568	6	ABU19714 Protein e
23	1914	63.8	566	6	ABU41362 Protein e
24	1907	63.6	568	6	ABU21842 Protein e

25	1897	63.3	568	6	ABU22314	Abu22314 Protein e
26	1832	61.1	566	6	ABU38819	Abu38819 Protein e
27	1828	61.0	566	7	ABO78141	ABO78141 Pseudomon
28	1826	60.9	572	6	ABU30286	Abu30286 Protein e
29	1825	60.9	567	6	ABU32083	Abu32083 Protein e
30	1821	60.7	625	7	ABO62699	ABO62699 Klebsiell
31	1818.5	60.6	569	7	ADF05744	Adf05744 Bacterial
32	1818.5	60.6	779	6	ABU41210	Abu41210 Protein e
33	1818	60.6	567	2	ADA34625	Ada34625 Acinetoba
34	1817	60.6	567	2	AAW37774	Aaw37774 Klebsiell
35	1817	60.6	567	3	AAV81825	Aav81825 Klebsiell
36	1815	60.5	566	6	ABU17040	Abu17040 Protein e
37	1811.5	60.4	568	6	AAK67379	Aar67379 P. mirabi
38	1809	60.3	568	6	ABU28899	Abu28899 Protein e
39	1809	60.3	568	6	ABU28920	Abu28920 Protein e
40	1809	60.3	568	7	ADC00113	Adc00113 Enterohae
41	1797	59.9	567	6	ABU40328	Abu40328 Protein e
42	1789.5	59.7	571	4	AAG81512	Aag81512 S. epider
43	1789.5	59.7	811	6	ABU42650	Abu42650 Protein e
44	1788.5	59.6	573	5	ABP40181	Abp40181 Staphyloc
45	1788.5	59.6	573	8	ADS07231	Ads07231 Staphyloc

ALIGNMENTS

RESULT 1
ADJ58239
ID ADJ58239 standard; protein; 568 AA.
XX
AC ADJ58239;
XX
DT 06-MAY-2004 (first entry)
XX
DE Urease subunit polypeptide Y.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX
OS Helicobacter felis.
XX
PN EPI176192-A2.
FN
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Kusters JG, Cattoli G;
XX
DR WPI; 2002-124384/17.
XX
DR N-PSDB; ADJ58237.
XX
PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
the diagnosis of Helicobacter felis infections and in the preparation of
vaccines.
XX
PS Claim 10; SEQ ID NO 3; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
Y subunit polypeptides and immunogenic fragments. The polypeptides are
used in the manufacture of vaccines against Helicobacter felis infections
and in diagnostic tests to detect antibodies against Helicobacter felis.
CC
CC Helicobacter felis is difficult to grow so it is more convenient to use
the expression products of the genes encoding the urease X and Y subunits
in the manufacture of vaccines. The present sequence represents an urease
Y subunit polypeptide of the invention.
XX
SQ Sequence 568 AA;
Query Match 100.0%; Score 2999; DB 5; Length 568;

Best Local Similarity 100.0%; Pred. No. 3.5e-253;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIRGMSGNSP 60
DB 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIRGMSGNSP 60

QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhMVVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhMVVGTEAL 120

QY 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDGNTATTTPGKNLHR 180
DB 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDGNTATTTPGKNLHR 180

QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

QY 241 DVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNLPSS 300
DB 241 DVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNLPSS 300

QY 301 PTIPYTIINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAW 360
DB 301 PTIPYTIINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAW 360

QY 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTH 420
DB 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTH 420

QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYREMF 480
DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYREMF 480

QY 481 HGKAKFDTISITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKIT 540
DB 481 HGKAKFDTISITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKIT 540

QY 541 TFEVFDGKLCSTKPTSQVPLAQRYTFF 568
DB 541 TFEVFDGKLCSTKPTSQVPLAQRYTFF 568

RESULT 2
ADJ58251
ID ADJ58251 standard; protein; 568 AA.

AC ADJ58251;
DT 06-MAY-2004 (first entry)
XX UreaseXY subunit #8.
DE immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
KW Helicobacter felis.
OS EP1176192-A2.
XX 30-JAN-2002.
PF 11-JUL-2001; 2001EP-00202666.
XX 17-JUL-2000; 2000EP-00202565.
PR (ALKU) AKZO NOBEL NV.
XX Kusters JG, Cattoli G;
PI WPI; 2002-124384/17.
XX N-PSDB; ADJ58249.
DR
XX

PT Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
XX vaccines.
XX Disclosure; SEQ ID NO 15; 76pp; English.
XX
CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
SQ Sequence 568 AA;

Query Match 99.1%; Score 2973; DB 5; Length 568;
Best Local Similarity 98.9%; Pred. No. 6.6e-251;
Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIRGMSGNSP 60
DB 1 MKMKQEVYNTYGTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIRGMSGNSP 60

QY 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhMVVGTEAL 120
DB 61 DENTLDLVITNAMIIDYTGKADIGIKNGKIHGIGKAGNKMDDGVSPhMVVGTEAL 120

QY 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDGNTATTTPGKNLHR 180
DB 121 AGEEMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTPGVDGNTATTTPGKNLHR 180

QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240
DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVAD 240

QY 241 DVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNLPSS 300
DB 241 DVQVCIHTDVTNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVITMAGELNLPSS 300

QY 301 PTIPYTIINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAW 360
DB 301 PTIPYTIINTVAEHLDMLTCHHLDKRIRDLQFSQSRIRPGSIAAEDVLHDMGVIAW 360

QY 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTH 420
DB 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTH 420

QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYREMF 480
DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPOPVYREMF 480

QY 481 HGKAKFDTISITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKIT 540
DB 481 HGKAKFDTISITFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKIT 540

QY 541 TFEVFDGKLCSTKPTSQVPLAQRYTFF 568
DB 541 TFEVFDGKLCSTKPTSQVPLAQRYTFF 568

RESULT 3
ADJ58242
ID ADJ58242 standard; protein; 568 AA.

XX ADJ58242;
AC ADJ58242;
DT 06-MAY-2004 (first entry)
XX UreaseXY subunit #2.
XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
KW

```
XX OS Helicobacter felis.
XX PN EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU ) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX PI WPI; 2002-124384/17.
XX DR N-PSDB; ADJ58240.
XX XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX PS Disclosure; SEQ ID NO 6; 76pp; English.
XX CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
XX SQ Sequence 568 AA;

Query Match 98.7%; Score 2960; DB 5; Length 568;
Best Local Similarity 98.6%; Pred. No. 9.1e-250;
Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMQSNP 60
DB 1 MKMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMQSNP 60
QY 61 DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
DB 61 DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
QY 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATTITPGKWNLHR 180
DB 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVFLGKGNSSSKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVFLGKGNSSSKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
QY 241 DVQVCITHTDVTNAGVDDTLNANMGRAIHAYHIEGAGGCHSPDVTMAGELNLPSTT 300
DB 241 DVQVCITHTDVTNAGVDDTLNANMGRAIHAYHIEGAGGCHSPDVTMAGELNLPSTT 300
QY 301 PTIPYINTVAEHLDMLTCHLDDKRIREDLQFSQRIIRPGSIAAEDVLHDMGVIAMTSS 360
DB 301 PTIPYINTVAEHLDMLTCHLDDKRIREDLQFSQRIIRPGSIAAEDVLHDI GVIAMTSS 360
QY 361 DSQAMGRAGEVIPRTWQTADRNKKEFGKLPEDGKNDNFRIKRYISKYINPALTHGVSE 420
DB 361 DSQAMGRAGEVIPRTWQTADRNKKEFGKLPEDGKNDNFRIKRYISKYINPALTHGVSE 420
QY 421 YIGSVBEKTIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYIREMFGH 480
DB 421 YIGSVBEKTIADLVVWNPAPFGVKPKIVIKGMVVFSEMGDSNASVPTPQVYIREMFGH 480
QY 481 HGKAKFDTSITTFVSKVAYENGVEKGLERKVLVKNCRNITKKDFKFNKTAHITVDPK 540
DB 481 HGKAKFDTSITTFVSKVAYENGVEKGLERKVLVKNCRNITKKDFKFNKTAHITVDPK 540

Db 481 HGKAKFDTSITTFVSKVAYENGVEKGLERKVLVKNCRNITKKDFKFNKTAHITVDPK 540
Qy 541 TFEVFDGKLGKTSKPTSQVPLAQRYTFF 568
Db 541 TFEVFDGKLGKTSKPTSQVPLAQRYTFF 568

RESULT 4
ADJ58248
ID ADJ58248 standard; protein; 568 AA.
XX AC ADJ58248;
XX XX 06-MAY-2004 (first entry)
XX XX UreaseXY subunit #6.
XX XX immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX OS Helicobacter felis.
XX PN EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU ) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX PI WPI; 2002-124384/17.
XX DR N-PSDB; ADJ58246.
XX XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX PS Disclosure; SEQ ID NO 12; 76pp; English.
XX CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
XX ureaseXY subunit of the invention.
XX SQ Sequence 568 AA;

Query Match 98.5%; Score 2953; DB 5; Length 568;
Best Local Similarity 98.2%; Pred. No. 3.7e-249;
Matches 558; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMQSNP 60
DB 1 MKMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTYGEEKLFGAGKTIREGMQSNP 60
QY 61 DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
DB 61 DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL 120
QY 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATTITPGKWNLHR 180
DB 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVFLGKGNSSSKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
DB 181 MLRAAEYSNMVFLGKGNSSSKQLVEQVGAIGFKLHEDWGTTPSAIDHCLSLVADEY 240
```

QY 241 DVQVCIHDTVNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDPVITMAGELNILPSSTT 300
DB 241 DVQVCIHDTVNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDPVITMAGELNILPSSTT 300
QY 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAEDVLHDMGVIAWTS 360
DB 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAEDVLHDMGVIAWTS 360
QY 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEKGKNDNFRIKRYISKYTTINPALTHGVSE 420
DB 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEGSADNDNFRIKRYISKYTTINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGVMVFPSEMGDSNASVPTPQPVYREMF 480
DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGVMVFPSEMGDSNASVPTPQPVYREMF 480
QY 481 HGKAKFDTISITFVSKVAYENGVEKLGLEKROVLVKNCRNITKKDPKFNKTAHITVDPK 540
DB 481 HGKAKFDTISITFVSKVAYENGVEKLGLEKROVLVKNCRNITKKDPKFNKTAHITVDPK 540
QY 541 TFEVFDGKLCSTKPSQVPLAQRITFF 568
DB 541 TFEVFDGKLCSTKPSQVPLAQRITFF 568

RESULT 5
ADJ58245
ID ADJ58245 standard; protein; 496 AA.
XX AC ADJ58245;
XX DT 06-MAY-2004 (first entry)
XX DE UreaseXY subunit #4.
XX KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis.
XX OS Helicobacter felis.
XX PN EP1176192-A2.
XX PD 30-JAN-2002.
XX PF 11-JUL-2001; 2001EP-00202666.
XX PR 17-JUL-2000; 2000EP-00202565.
XX PA (ALKU) AKZO NOBEL NV.
XX PI Kusters JG, Cattoli G;
XX DR WPI: 2002-124384/17.
XX DR N-PSDB; ADJ58243.

XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT the diagnosis of Helicobacter felis infections and in the preparation of
PT vaccines.
XX Disclosure; SEQ ID NO 9; 76pp; English.
XX CC The present invention relates to a novel Helicobacter felis urease X and
CC Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC used in the manufacture of vaccines against Helicobacter felis infections
CC and in diagnostic tests to detect antibodies against Helicobacter felis.
CC Helicobacter felis is difficult to grow so it is more convenient to use
CC the expression products of the genes encoding the urease X and Y subunits
CC in the manufacture of vaccines. The present sequence represents an
CC ureaseXY subunit of the invention.
XX
XX Sequence 496 AA;
Query Match 86.3%; Score 2588; DB 5; Length 496;
Best Local Similarity 99.4%; Pred. No. 2.8e-217;

Matches 489; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMQSNP 60
DB 1 MKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYTGELKFGAGKTIREGMQSNP 60
QY 61 DENTLDLVTNAMIIDYTGIIKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTAL 120
DB 61 DENTLDLVTNAMIIDYTGIIKADIGIKNGKIHGIGKAGNKMDQGVSPHVVGVGTAL 120
QY 121 AGEEMIITAGGIDSHFHSFPQPTALANGVTTMFGGTGPVDTGTTNATTITPGKWNLHR 180
DB 121 AGEEMIITAGGIDSHFHSFPQPTALANGVTTMFGGTGPVDTGTTNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEY 240
DB 181 MLRAAEYSNMVFLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADEY 240
QY 241 DVQVCIHDTVNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDPVITMAGELNILPSSTT 300
DB 241 DVQVCIHDTVNEAGVDDTLNAMGRAIHAYHIEGAGGHSFDPVITMAGELNILPSSTT 300
QY 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAEDVLHDMGVIAWTS 360
DB 301 PTPIPTINTVAEHLDMLTCHLDKRIREDLOFSQSRIRPGSIAEDVLHDMGVIAWTS 360
QY 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEKGKNDNFRIKRYISKYTTINPALTHGVSE 420
DB 361 DSQAMGRAGEVIPRTWQTADKNKCFGLPEKGKNDNFRIKRYISKYTTINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGVMVFPSEMGDSNASVPTPQPVYREMF 480
DB 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGVMVFPSEMGDSNASVPTPQPVYREMF 480
QY 481 HGKAKFDTISITF 492
DB 481 HGKAKFDTISITF 492
RESULT 6
ADQ37851
ID ADQ37851 standard; protein; 569 AA.
XX AC ADQ37851;
XX DT 07-OCT-2004 (first entry)
XX DE H. bizzozeronii ureB polypeptide.
XX KW Urease; urease gene cluster; urease structural gene;
KW urease accessory gene; ureB; Helicobacter bizzozeronii infection;
KW antibacterial; enzyme.
XX OS Helicobacter bizzozeronii.
XX PN US2004142343-A1.
XX PD 22-JUL-2004.
XX PF 12-AUG-2003; 2003US-00639273.
XX PR 16-AUG-2002; 2002US-0404337P.
XX PA (CHAN/) CHANG Y.
XX PA (SIMP/) SIMPSON K W.
XX PA (ZHUJ/) ZHU J.
XX PI Chang Y, Simpson KW, Zhu J;
XX DR WPI: 2004-533502/51.
XX DR N-PSDB; ADQ37850.
XX PT Novel isolated nucleic acid molecule having urease gene cluster, and

Dd 123 GGLIVTAGGIDTHIFISPOOIPITAFASGVVTMTIGGGTGPADGNTNATITTPGRNRLKWM 182
Qy 182 LRAAEYSNNVFLGKGNSSKKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Dd 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIQFKIHEDWGTTPSAINHALDVADKYD 242
Qy 242 VQVCITHTDVTNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTTP 301
Dd 243 VQVAIHTDVTNEAGCVEDTMAAIAGRMTHTFTEGAGGSHAPDIIKVAGEHNILPASTNP 302
Qy 302 TTPYTINTVAEHLDMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLDHMGVIAMTSSD 361
Dd 303 TTPFTVTNEAEHMDMLVCHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFISITSD 362
Qy 362 SQAMGRAGEVIRPTWOTADKNKEFGKLPEDGKNDNFRIKRYISKYTINPALTHTGVSY 421
Dd 363 SQAMGRVGEVITRTWQTADKNKEFGRLKEEGKNDNFRIKRYLSKYTINPAIAHGISY 422
Qy 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTQPQVYVREMFHGH 481
Dd 423 VGSVEVGKVDLVWSPAPFFGVKPNMIKGGFIALSQMGDANAS IPTQPQVYVREMFPAH 482
Qy 482 GKAKEDTSITFSKVAYENGVEKGLERQVLPVKCRNITKDFKFNDKTAKITVDPKT 541
Dd 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLPVKCRNITKDMQFNDTTAHIEVNPET 542
Qy 542 FEVFDGKLTCKTSKPTSOVPLAORYTFF 568
Dd 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 11
ADU05536
ID ADU05536 standard; protein; 569 AA.
XX
AC ADU05536;
XX
DT 27-JAN-2005 (first entry)
XX
DE H. pylori antigenic protein HP0072.
XX
KW antibacterial; antigenic; H. pylori infection.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT Region 533..567
FT /note= "Immunogenic region. Region specifically claimed
FT in claim 14"
XX
PN WO2004094467-A2.
XX
PD 04-NOV-2004.
XX
PF 22-APR-2004; 2004WO-BP004255.
XX
PR 22-APR-2003; 2003EP-00450097.
XX
PA (INTE-) INTERCELL AG.
XX
PI Meinke A, Min Bui D, Nagy E, Henics T;
XX
XX WPI; 2004-775908/76.
DR N-PSDB; ADU05358.
XX
XX New hyperimmune serum reactive antigens from Helicobacter pylori, and
PT encoding nucleic acid molecules, useful for diagnosing, preventing or
PT treating H. pylori infections.
XX
PS Claim 14; SEQ ID NO 184; 176pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC hyperimmune serum reactive antigen or its fragment. The composition

CC (including the nucleic acid molecule, hyperimmune serum-reactive antigen
CC or antibody) is useful for manufacturing a medicament or pharmaceutical
CC preparation (e.g. a vaccine) for treating or preventing H. pylori
CC infections. The antigen or its fragment may also be used for isolating,
CC purifying and/or identifying an interaction partner of the hyperimmune
CC serum reactive antigen or fragment; for generating a peptide binding to
CC the hyperimmune serum reactive antigen or fragment, where the peptide is
CC selected from aptamers and spiegelmers; or for manufacturing a functional
CC ribonucleic acid selected from ribozymes, antisense nucleic acids and
CC siRNA. The present sequence represents the amino acid sequence of an H.
CC pylori antigenic protein.
XX
SQ Sequence 569 AA;
Query Match 76.6%; Score 2297; DB 8; Length 569;
Best Local Similarity 73.4%; Pred. No. 1.1e-191;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
Qy 2 KMKKQYVNTYGTGDKVRLGDTDLWAEVHDYTYGEELKFGAGKTIREGMGQSNPSD 61
Dd 3 KISRKEYVSMGYPTTGDKVRLGDTDLAEVHDYTYGEELKFGGKTLREGMSQSNPS 62
Qy 62 ENTLDLVITNAMIIDVTGIYKADIGIKNGIKHIGKAGNKDMQGVSPHVVGVTEALA 121
Dd 63 KEELDLIITNALIVDTYGIYKADIGIKGKIAGIKGKNDKMDQGVKNLNVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQOPTALANGVTMMFGGTPGVDGNTATITTPGKNLHRM 181
Dd 123 GEGLIVTAGGIDTHIFISPOOIPITAFASGVVTMTIGGGTGPADGNTNATITTPGRNRLKWM 182
Qy 182 LRAAEYSNNVFLGKGNSSKKQLVEQVAGAIQFKLHEDWGTTPSAIDHCLSVADYD 241
Dd 183 LRAAEYSNNLGLAKGNASNDASLADQIEAGAIQFKIHEDWGTTPSAINHALDVADKYD 242
Qy 242 VQVCITHTDVTNEAGYVDDTLNMGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTTP 301
Dd 243 VQVAIHTDVTNEAGCVEDTMAAIAGRMTHTFTEGAGGSHAPDIIKVAGEHNILPASTNP 302
Qy 302 TTPYTINTVAEHLDMTCHLDKRIREDLQFSQSRIRPGSIAAEDVLDHMGVIAMTSSD 361
Dd 303 TTPFTVTNEAEHMDMLVCHLDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFISITSD 362
Qy 362 SQAMGRAGEVIRPTWOTADKNKEFGKLPEDGKNDNFRIKRYISKYTINPALTHTGVSEY 421
Dd 363 SQAMGRVGEVITRTWQTADKNKEFGRLKEEGKNDNFRIKRYLSKYTINPAIAHGISY 422
Qy 422 IGSVEEGKIADLVVWNPAPFFGVKPKIVIKGGMVVFSEMGDSNASVPTQPQVYVREMFHGH 481
Dd 423 VGSVEVGKVDLVWSPAPFFGVKPNMIKGGFIALSQMGDANAS IPTQPQVYVREMFPAH 482
Qy 482 GKAKEDTSITFSKVAYENGVEKGLERQVLPVKCRNITKDFKFNDKTAKITVDPKT 541
Dd 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLPVKCRNITKDMQFNDTTAHIEVNPET 542
Qy 542 FEVFDGKLTCKTSKPTSOVPLAORYTFF 568
Dd 543 YHVFVDGKEVTSKPANKVSLAQFSIF 569

RESULT 12
ADS09177
ID ADS09177 standard; protein; 568 AA.
XX
AC ADS09177;
XX
DT 16-DEC-2004 (first entry)
XX
DE H. pylori urease beta subunit.
XX
KW Urease; beta subunit; UreB; gastric ulcer; peptic ulcer; gastritis;
KW stomach cancer; vaccine; antibody; immune reaction.
XX

QY 122 GEGMIITAGGIDSHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATITTPGKNLHRM 181
Db 123 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTMTMIGGGTGPADGTTNATITTPGRRNLKWM 182
QY 182 LRAAEYSNMVGLGKGNSSKQLVEQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 241
Db 183 LRAAEYSNMVGLGKGNSSKQLVEQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 242
QY 242 VQVCIHTDVNAGYVDDTLNMGNGRAIHAYHIEGAGGHSPPDVTMAGELNLPSTTP 301
Db 243 VQVAIHTDNLNAGCVEDTMAAIGRTMTFTHTGAGGGHAPDIIKVAGEHNILPASTNP 302
QY 302 TTPYTINTVAEHLDMVCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 361
Db 303 TTPFTVNTAEHMDMLVCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 362
QY 362 SOAMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 363 SOAMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 422
QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 481
Db 423 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 482
QY 482 GKAKPDTSTTFVSKVAYENGVEKLGRLQVLPVKNCRNITKKDFKFNDAKTAKITVDPKT 541
Db 483 GKAKYDRNITFVSQAAYDKGKEELGLERQVLPVKNCRNITKKDMQFNDTTAHIEVNPET 542
QY 542 FEVFDGKLCSTKPTSQVPLAQRITFF 568
Db 543 YHVFVDGKEVTSKPNKVSQAQLFSIF 569

RESULT 14

ABE55116

ID ABE55116 standard; protein; 570 AA.

AC ABE55116;

DT 22-SEP-2005 (first entry)

DE Helicobacter pylori Urase B protein fragment SEQ ID 5.

KW pHUR3; vaccine; helicobacter pylori infection; urease;
KW immune stimulation; antibacterial; immunostimulant; antiulcer;
KW cytotstatic; antiinflammatory; gastrointestinal-gen.; gastritis;
KW peptic ulcer; carcinoma.

OS Helicobacter pylori.

PN WO200132014-A2.

XX 10-MAY-2001.

XX 01-NOV-2000; 2000WO-US030191.

XX 01-NOV-1999; 99US-00431705.

XX (ORAV-) ORAVAX INC.

PI Kleanthous H, Londono-Arcila P, Freeman D;

XX WPI; 2001-343379/36.

DR N-PSDB; ABE55112, ABE55130.

XX Inducing an immune response against Helicobacter in mammals, useful for
PT treating Helicobacter induced gastroduodenal diseases.
XX Disclosure; SEQ ID NO 5; 63pp; English.

XX The invention relates to inducing an immune response against Helicobacter

CC in a mammal, comprising mucosally administering to the mammal an
CC attenuated Salmonella vector containing a nucleic acid molecule encoding

CC a Helicobacter antigen, and parenterally administering to the mammal a
CC Helicobacter antigen. Also included is an attenuated Salmonella vector
CC comprising a nucleic acid molecule encoding a Helicobacter antigen. The
CC Helicobacter antigen is a urease, a urease subunit, or its immunogenic
CC fragment (encoded by the ureA and ureB genes). The mammal is at risk of
CC having but does not have Helicobacter infection or has a Helicobacter
CC infection. The attenuated Salmonella vector further comprises an htrA or
CC ntrB promoter. The vector can be used in inducing an immune response
CC against Helicobacter in a mammal. The vector can be used to treat
CC Helicobacter infection. The vector and the method can be used to treat
CC Helicobacter induced gastroduodenal diseases, including acute, chronic or
CC atrophic gastritis, peptic ulcer diseases (e.g. gastric or duodenal
CC ulcers), and gastric carcinoma. Attenuated Salmonella vector pHUR3 was
CC constructed expressing UreA and UreB and including an E. coli Amp^r gene.
CC The present sequence represents a protein or peptide encoded by the pHUR3
CC plasmid. NOTE: pHUR3 is represented by both ABE55112 and ABE55130, the
CC peptides expressed by pHUR3 are shown in the sequence listing to be split
CC between ABE55112 and ABE55130, yet Figure 4 shows all the peptides being
CC expressed by the one sequence, therefore all encoded peptides are cross-
CC referenced to both sequences.

XX SQ Sequence 570 AA;

Query Match 76.1%; Score 2281; DB 4; Length 570;

Best Local Similarity 73.0%; Pred. No. 2.7e-190;

Matches 414; Conservative 76; Mismatches 77; Indels 0; Gaps 0;

QY 2 KMKKQEVNNTYPTGDKVRLGDDTLWAEVDEHYTTYGEELKFGAGKTIREMGOSNSPD 61

Db 4 KISRKQEVNNTYPTGDKVRLGDDTLWAEVDEHYTTYGEELKFGAGKTIREMGOSNSPD 63

QY 62 ENTLDLVITNAMIIDYTYGIYKADIGIKNGKIHGICAGNKMDDQGVSPHVVGVTEALA 121

Db 64 KEELDLIITNALIVDYTYGIYKADIGIKNGKIHGICAGNKMDDQGVSPHVVGVTEALA 123

QY 122 GEGMIITAGGIDSHFLSPQFPPTALANGVTTMFGGGTGPVDGTTNATITTPGKNLHRM 181

Db 124 GEGLIIVTAGGIDTHIFISPOQIPTAFASGVTMTMIGGGTGPADGTTNATITTPGRRNLKWM 183

QY 182 LRAAEYSNMVGLGKGNSSKQLVEQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 241

Db 184 LRAAEYSNMVGLGKGNSSKQLVEQVEAGAIGFKLHEDMGTTPSAIDHCLSVADVD 242

QY 242 VQVCIHTDVNAGYVDDTLNMGNGRAIHAYHIEGAGGHSPPDVTMAGELNLPSTTP 301

Db 244 VQVAIHTDNLNAGCVEDTMAAIGRTMTFTHTGAGGGHAPDIIKVAGEHNILPASTNP 303

QY 302 TTPYTINTVAEHLDMVCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 361

Db 304 TTPFTVNTAEHMDMLVCHLDRKIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTSDD 363

QY 362 SOAMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421

Db 364 SOAMGRAGEVIRPTWQADKNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 423

QY 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 481

Db 424 VGSVEVGKADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYREMFHGH 483

QY 482 GKAKPDTSTTFVSKVAYENGVEKLGRLQVLPVKNCRNITKKDFKFNDAKTAKITVDPKT 541

Db 484 GKAKYDRNITFVSQAAYDKGKEELGLERQVLPVKNCRNITKKDMQFNDTTAHIEVNPET 543

QY 542 FEVFDGKLCSTKPTSQVPLAQRITFF 568

Db 544 YHVFVDGKEVTSKPNKVSQAQLFSIF 570

RESULT 15

AAR74337

ID AAR74337 standard; protein; 569 AA.

XX AAR74337;

AC AAR74337;

Qy 362 SQAMGRAGEVTPRTWOTADKNKKEFGKLPEDGKNDNFRISKYISKYTINPALTHGVSEY 421
Db 363 SQAMGRVGEVITRTWOTADKNKKEFGRLKEBKGDNDNFRISKYISKYTINPGIAHGISDY 422
Qy 422 IGSVEEGKIADLVVNPAFFGVGPKIVIKGGMVWFSEMGDSNASVPTQPVPVYRPMFGHH 481
Db 423 VGSVEVGKYADLVLMSPAFFGKIPNMIKGGFIALSQMGDANASITPTQPVPVYRPMFGHH 482
Qy 482 GKAKFDTSITFVSKVAYENGVEKLGLEQLERQVLPVKNCRNITKQDFKNDKTAKITVDPKT 541
Db 483 GKAKFDNITFVSQAAYKAGIKBELGLDRAAPPVKNCRNITKDKLKFNDDVTAHIDVNPET 542
Qy 542 FEVFDGKLTCKTSKPTSQVPLAQRTPFF 568
Db 543 YKVKVDGKEVTSKAADELSLAQLNLF 569

Search completed: November 28, 2005, 08:14:37
Job time : 188.28 secs

XX 25-MAR-2003 (revised)
DT 02-NOV-1995 (first entry)
XX Helicobacter felis urease UreB.
XX Urease; UreA; UreB; vaccine; Helicobacter infection; UreI;
KW heat shock protein.
XX Helicobacter felis.
OS Helicobacter felis.
XX WO9514093-A1.
PN
PD
XX 26-MAY-1995.
XX 19-NOV-1993; 93WO-EF003259.
PF 19-NOV-1993; 93WO-EF003259.
PR (INSP) INST PASTEUR.
PA (INRM) INST NAT SANTE & RECH MEDICALE.
XX
PI Labigne A, Suerbaum S, Ferrero R;
XX
DR WPI; 1995-200383/26.
DR N-PSDB; AAQ90180.
XX
PT Immunogenic composition against Helicobacter infection - also gene
PT fragment(s) and protein(s) from Helicobacter urease gene cluster and heat
PT shock protein(s).
XX
PS Claim 10; Fig 3; 128pp; English.
XX
CC The urease UreB protein is a component of a novel immunogenic composition
CC capable of inducing protective antibodies against Helicobacter infection.
CC The composition may include the UreA protein (AAR74336), a urease-
CC associated heat shock protein (AAR74338-39) or the UreI protein
CC (AAR74340). The composition is used to prepare a vaccine for humans or
CC animals, especially against H. pylori and H. felis. Antibodies against
CC the proteins may be used for treating Helicobacter infection, and
CC primers/probes to the DNA sequence may be used for detection of
CC Helicobacter infection. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 569 AA;
Query Match 74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 4.7e-187;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;
Qy 2 KMKKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGBELKEGAGKTIREGMGQSNSPD 61
Db 3 KISRKEYSMYGPTTGDVRLGDTDLILEVHDCTTYGEEIKFGGKTRDGMSTNSPS 62
Qy 62 ENTDLVITNAMIIDYTIYKADIGIKNGKIHGIGKAGNKMDQGVSPHMVGVGTEALA 121
Db 63 SYELDLVLNLIYDVTGYIKADIGIKGKAGNKMDQGVDDNLCVGPATEALA 122
Qy 122 GEGMIITAGDISHTHFLSPQPPPTALANGVTMTFGGTPVDGTNATTITPKWNLHRM 181
Db 123 AEGLIVTAGDITHIFISPPQIPTAFASGVTMTIGGTPADGTNATTITPGRANLKS 182
Qy 182 LRAAEYSNNVFLGKGNSSSKQLVEQVEAGIAGFKLHEDGTTSPADHCLSVAD EYD 241
Db 183 LRAAEYAMNLGFLAKGNVSYEPLSDQIEAGIAGFKIHEDMGSTPAAIHCLNVADEYD 242
Qy 242 VQVCHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHS PDVITMAGELNLPSTTP 301
Db 243 VQVAHTDTLNEAGCVEDLEATAGTTHFTTEGAGGGHAPDVIKXAGEFNILPASTNP 302
Qy 302 TIPTYINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TIPFTKNTAEHMDMLMVCHLDKSIKEDVQPADSRIRPQTIAAEDQLHDMGIFSISSD 362

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:11:38 ; Search time 160.242 Seconds
(without alignments)
1481.056 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKMKQEVNTYGTGDKV.....KLCTSKPTSQVPLAQRYTF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap:
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap:
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap:
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pap:
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap:
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match %	Length	DB ID	Description	
1	2999	100.0	568	3	US-09-904-994B-3	Sequence 3, Appli
2	2973	99.1	568	3	US-09-904-994B-15	Sequence 15, Appl
3	2960	98.7	568	3	US-09-904-994B-6	Sequence 6, Appli
4	2953	98.5	568	3	US-09-904-994B-12	Sequence 12, Appl
5	2588	86.3	496	3	US-09-904-994B-9	Sequence 9, Appli
6	2326	77.6	569	4	US-10-639-273-5	Sequence 5, Appli
7	2298.5	76.6	568	4	US-10-639-273-38	Sequence 38, Appl
8	2297	76.6	569	3	US-09-895-913A-252	Sequence 252, App
9	2297	76.6	569	4	US-10-282-122A-58611	Sequence 58611, A
10	2297	76.6	569	4	US-10-335-977-8611	Sequence 8611, Ap
11	2297	76.6	569	4	US-10-639-273-40	Sequence 40, Appl
12	2244	74.8	569	4	US-10-639-273-39	Sequence 39, Appl
13	2234	74.5	750	3	US-09-402-100-2	Sequence 2, Appli
14	2107	70.3	569	4	US-10-639-273-41	Sequence 41, Appl
15	1922	64.1	568	4	US-10-282-122A-47638	Sequence 47638, A
16	1914	63.8	566	4	US-10-282-122A-69286	Sequence 69286, A
17	1907	63.6	568	4	US-10-282-122A-49766	Sequence 49766, A
18	1897	63.3	568	4	US-10-282-122A-50238	Sequence 50238, A
19	1832	61.1	566	4	US-10-282-122A-66743	Sequence 66743, A
20	1826	60.9	572	4	US-10-282-122A-58210	Sequence 58210, A
21	1825	60.9	567	4	US-10-282-122A-60007	Sequence 60007, A
22	1818.5	60.6	779	4	US-10-282-122A-69134	Sequence 69134, A
23	1815	60.5	566	4	US-10-282-122A-44964	Sequence 44964, A
24	1809	60.3	568	4	US-10-282-122A-56823	Sequence 56823, A
25	1809	60.3	568	4	US-10-282-122A-56844	Sequence 56844, A
26	1807	60.3	837	4	US-10-424-599-254635	Sequence 254635, A
27	1797	59.9	567	4	US-10-282-122A-68252	Sequence 68252, A

RESULT 1

US-09-904-994B-3
; Sequence 3, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO NO. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
; US-09-904-994B-3

Query Match 100.0%; Score 2999; DB 3; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.4e-258;
Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGEBELKFGAGKTIREGMGQSNP	60
Db	1	MMKMKQEVNTYGTGDKVRLGDTDLWAEVEHDYTTYGEBELKFGAGKTIREGMGQSNP	60
Qy	61	DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL	120
Db	61	DENTLDLVIITNAMIIDYTYGKADIGIKNGKIHGKAGNKMDDGVSPHVMVGVGTEAL	120
Qy	121	AGGMIITAGIDISHTHFLSPQQPPTALANGVTMTMFGGGTGPVDCGTNATTITPGKNLHR	180
Db	121	AGGMIITAGIDISHTHFLSPQQPPTALANGVTMTMFGGGTGPVDCGTNATTITPGKNLHR	180
Qy	181	MLRAAEYSNMVGLFGKNSSSKKQLVEQVEAGAIGFKLHEDWGTPPSAIDHCLSVADY	240
Db	181	MLRAAEYSNMVGLFGKNSSSKKQLVEQVEAGAIGFKLHEDWGTPPSAIDHCLSVADY	240
Qy	241	DVQVCIHDTVNEAGYVDDTLNANGRAIHAIHYHIEGAGGHSPPVITWAGELNIIIPSSYT	300
Db	241	DVQVCIHDTVNEAGYVDDTLNANGRAIHAIHYHIEGAGGHSPPVITWAGELNIIIPSSYT	300
Qy	301	PTTPYTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSS	360
Db	301	PTTPYTINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSS	360
Qy	361	DSQAMGRAGEVIPRTWTQADKNKKEFGKLPDGDKNNDNFRIKRYISKYITINPALTGHVSE	420
Db	361	DSQAMGRAGEVIPRTWTQADKNKKEFGKLPDGDKNNDNFRIKRYISKYITINPALTGHVSE	420

Qy 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRYYTF 568
Db 541 TFEVFDGKCLCTSKPTSOVPLAQRYYTF 568
RESULT 2
US-09-904-994B-15
; Sequence 15, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-15
Query Match 99.1%; Score 2973; DB 3; Length 568;
Best Local Similarity 98.9%; Pred. No. 2.9e-256;
Matches 562; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKMKQEVYNTYGPPTGDKVRLGDTDLMAVEHDYTTYGEELKFGAGKTI REGMGQSNP 60
Db 1 MKMKQEVYNTYGPPTGDKVRLGDTDLMAVEHDYTTYGEELKFGAGKTI REGMGQSNP 60
Qy 61 DENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Db 61 DENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Qy 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPGVDGTNATTITPGKNLHR 180
Db 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPGVDGTNATTITPGKNLHR 180
Qy 181 MLRAAEEYSNMVGFGLKGNSSSKQLVEQVEAGAIGFKLHEDWGTTTSAIDHCLSVADY 240
Db 181 MLRAAEEYSNMVGFGLKGNSSSKQLVEQVEAGAIGFKLHEDWGTTTSAIDHCLSVADY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVI TMAGELNLPSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVI TMAGELNLPSTT 300
Qy 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTS 360
Db 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTGVSE 420
Qy 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRYYTF 568
Db 541 TFEVFDGKCLCTSKPTSOVPLAQRYYTF 568

Db 541 TFEVFDGKCLCTSKPTSEVPPLAQRYYTF 568
RESULT 3
US-09-904-994B-6
; Sequence 6, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-6
Query Match 98.7%; Score 2960; DB 3; Length 568;
Best Local Similarity 98.6%; Pred. No. 4.2e-255;
Matches 560; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MKMKQEVYNTYGPPTGDKVRLGDTDLMAVEHDYTTYGEELKFGAGKTI REGMGQSNP 60
Db 1 MKMKQEVYNTYGPPTGDKVRLGDTDLMAVEHDYTTYGEELKFGAGKTI REGMGQSNP 60
Qy 61 DENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Db 61 DENTLDLVTNAMIIDYTGIIYKADIGIKNGKIHGIGKAGNKMDQGVSPHVMVGVGTEAL 120
Qy 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPGVDGTNATTITPGKNLHR 180
Db 121 AGEGLMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPGVDGTNATTITPGKNLHR 180
Qy 181 MLRAAEEYSNMVGFGLKGNSSSKQLVEQVEAGAIGFKLHEDWGTTTSAIDHCLSVADY 240
Db 181 MLRAAEEYSNMVGFGLKGNSSSKQLVEQVEAGAIGFKLHEDWGTTTSAIDHCLSVADY 240
Qy 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVI TMAGELNLPSTT 300
Db 241 DVQVCIHDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVI TMAGELNLPSTT 300
Qy 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTS 360
Db 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAWTS 360
Qy 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTGVSE 420
Db 361 DSQAMGRAGEVIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTIINPALTGVSE 420
Qy 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Db 421 YIGSVEEGKIADLVVWNPAPFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVVYREMFH 480
Qy 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Db 481 HGKAKFDSITFVSKVAYENGVEKGLERQVLVVKNCRNITKKDFKNDKTAKITVDPK 540
Qy 541 TFEVFDGKCLCTSKPTSOVPLAQRYYTF 568
Db 541 TFEVFDGKCLCTSKPASEVPLAQRYYTF 568
RESULT 4
US-09-904-994B-12
; Sequence 12, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine

; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-12

```
Query Match      98.5%; Score 2953; DB 3; Length 568;
Best Local Similarity 98.2%; Pred. No. 1.8e-254;
Matches 558; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKMKQEVVNTYGTPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMSGNSP 60
DB 1 MKMKQEVVNTYGTPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMSGNSP 60
QY 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGKMDQGVSPHVVGVGTEAL 120
DB 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGKMDQGVSPHVVGVGTEAL 120
QY 121 AGEEMIITAGGIDSHFLSPQFPPTALANGVTTMFGGTGPVDTGNATTITPGKWNLHR 180
DB 121 AGEEMIITAGGIDSHFLSPQFPPTALANGVTTMFGGTGPVDTGNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADY 240
DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADY 240
QY 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSSTT 300
DB 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSSTT 300
QY 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
DB 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
QY 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
DB 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYREMFGH 480
DB 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYREMFGH 480
QY 481 HGKAKFDTSIPTVSKVAYENGVEKLGRLQVLPVKNCRNITKDKPFNDKTAKITVDPK 540
DB 481 HGKAKFDTSIPTVSKVAYENGVEKLGRLQVLPVKNCRNITKDKPFNDKTAKITVDPK 540
QY 541 TFEVFDGKLTCKTSKPTSQVPLAQRVTFP 568
DB 541 TFEVFDGKLTCKTSKPTSQVPLAQRVTFP 568
```

RESULT 5
US-09-904-994B-9
; Sequence 9, Application US/09904994B
; Publication No. US20040005325A1
; GENERAL INFORMATION:
; APPLICANT: AKZO No. US20040005325A1el N.V.
; TITLE OF INVENTION: Helicobacter vaccine
; FILE REFERENCE: Degoodesequences
; CURRENT APPLICATION NUMBER: US/09/904,994B
; CURRENT FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Helicobacter felis
US-09-904-994B-9

```
Query Match      86.3%; Score 2588; DB 3; Length 496;
Best Local Similarity 99.4%; Pred. No. 5.8e-222;
Matches 489; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMKQEVVNTYGTPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMSGNSP 60
DB 1 MKMKQEVVNTYGTPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMSGNSP 60
QY 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGKMDQGVSPHVVGVGTEAL 120
DB 61 DENTLDLVTNAMIIDYTGKADIGIKNGKIHGIGKAGKMDQGVSPHVVGVGTEAL 120
QY 121 AGEEMIITAGGIDSHFLSPQFPPTALANGVTTMFGGTGPVDTGNATTITPGKWNLHR 180
DB 121 AGEEMIITAGGIDSHFLSPQFPPTALANGVTTMFGGTGPVDTGNATTITPGKWNLHR 180
QY 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADY 240
DB 181 MLRAAEYSNMVGFGLGKGNSSKKQLVEQVEAGAIGFKLHEDWGTTPSAIDHCLSVADY 240
QY 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSSTT 300
DB 241 DVQVCIHTDTVNEAGYVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTIMAGELNIPSSTT 300
QY 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
DB 301 PTIPYTIINTVAEHLDMLTCHLDKRIREDLQFSQSRIRPGSIAEDVLHDMGVIAWTS 360
QY 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
DB 361 DSQAMGRAGEVTPRTWQTADKNKEFGKLPEDGKNDNFRIKRYISKYTIINPALTHGVSE 420
QY 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYREMFGH 480
DB 421 YIGSVEEGKIADLVVWNPFAFGVKPKIVIKGMVVFSEMGDSNASVPTPQPVYREMFGH 480
QY 481 HGKAKFDTSIPT 492
DB 481 HGKAKFDTSIPT 492
```

RESULT 6
US-10-639-273-5
; Sequence 5, Application US/10639273
; Publication No. US20040142343A1
; GENERAL INFORMATION:
; APPLICANT: Chang, Yung-Fu
; APPLICANT: Simpson, Kenneth W
; APPLICANT: Zhu, Jiaguan
; TITLE OF INVENTION: HELICOBACTER BIZZOZERONII UREASE GENES AND THEIR USES
; TITLE OF INVENTION: IN DIAGNOSTIC AND TREATMENT METHODS
; FILE REFERENCE: 19603/3881
; CURRENT APPLICATION NUMBER: US/10/639,273
; CURRENT FILING DATE: 2003-08-12
; PRIOR APPLICATION NUMBER: 60/404,337
; PRIOR FILING DATE: 2002-08-16
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter bizzozeronii
US-10-639-273-5

```
Query Match      77.6%; Score 2326; DB 4; Length 569;
Best Local Similarity 74.1%; Pred. No. 1.8e-198;
Matches 420; Conservative 75; Mismatches 72; Indels 0; Gaps 0;

QY 2 MKMKQEVVNTYGTPTGDKVRLGDTDLWAEVEHDYTTYGEELKFGAGKTIREGMSGNSPD 61
DB 3 KISRKEYVSMYGTPTGDKVRLGDTDLILEVHDCTTYGEEKIKFGGGKTIRDMAGTNSPS 62
```



```
Qy 182 LRAAEYSNMVGLFGKGNSSKKQLVEQVEAGAIGFKLHEDWCTTPSAIDHCLSVAD EYD 241
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 183 LRAAEYSNMVGLFLAKGNASNDASLADQIEAGAIGFKIHEDWCTTPSAINHALDVADKYD 242
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 242 VQVCIHTDVTNBAAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPVITWAGELNLPSTTP 301
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 243 VQVAIHDTLNEAGCVEDTAAIAGRTMHTFHTGAGGGHAPDIIKVAAGEHNILPASTNP 302
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 302 TTPYINTVAEHLDMTCHLDRREDLQFSQSRIRGCSIAAEDVLHDMGVIAWTSDD 361
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 303 TTPFTVNTAEHMDMLVCHLDRSFKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 362 SQAMGRAGVEIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTNPALTHGVSEY 421
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEEKDNDNFRIKRYLSKYTNPAIAHGISEY 422
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGSDSNASVPTPQVYVREMFHGH 481
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 423 VGSVEVGKVLVLVWSPAFFGKPNMIKGGFIALSQMGDANASIPTPQVYVREMFHAH 482
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLVKNCRNITKKDMQFNDTTAHIEVNPET 542
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 542 FEVFDGKLCSTKPTSOVPLAQRVTF 568
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

RESULT 9

```
US-10-282-122A-58611
; Sequence 58611, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58611
; LENGTH: 569
; TYPE: PRF
; ORGANISM: Helicobacter pylori
; US-10-282-122A-58611
```

```
Query Match 76.6%; Score 2297; DB 4; Length 569;
Best Local Similarity 73.4%; Pred. No. 7.2e-196;
Matches 416; Conservative 76; Mismatches 75; Indels 0; Gaps 0;
```

```
Qy 2 KMKKQEVNVTGPTKGDKVRIGDLDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNSPD 61
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 KISRKEYVSMGPTTGDKVRIGDLDLMAEVEHDYTYGEEELKFGAGKTIREGMGQSNPNFS 62
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 62 ENTLDLVTNAMIIDYTIYKADIGIKNGKIHGKAGNKMDDQGVSPHMVVGVTGEALA 121
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 63 KEELDLITNALIVDYTIYKADIGIKDGKIAGIKGNGKMDQGVKNNLSVGPATEALA 122
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 122 GEGMIITAGGIDSHTEFLSPQFPFTALANGVTTFMFGGCTGPDVGTNATTTIPGKNLHRM 181
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 123 GEGLIVTAGGIDTHIHFIHQIIFAFASGVTTMIGGCTGPDADGTNATTTIPGRNLKWM 182
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 182 LRAAEYSNMVGLFGKGNSSKKQLVEQVEAGAIGFKLHEDWCTTPSAIDHCLSVAD EYD 241
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 183 LRAAEYSNMVGLFLAKGNASNDASLADQIEAGAIGFKIHEDWCTTPSAINHALDVADKYD 242
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 242 VQVCIHTDVTNBAAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPVITWAGELNLPSTTP 301
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 243 VQVAIHDTLNEAGCVEDTAAIAGRTMHTFHTGAGGGHAPDIIKVAAGEHNILPASTNP 302
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 302 TTPYINTVAEHLDMTCHLDRREDLQFSQSRIRGCSIAAEDVLHDMGVIAWTSDD 361
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 303 TTPFTVNTAEHMDMLVCHLDRSFKEDVQFADSRIRPQTIAAEDTLHDMGIFSITSSD 362
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 362 SQAMGRAGVEIPRTWQTADKNKKEFGKLPEDGKNDNFRIKRYISKYTNPALTHGVSEY 421
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 363 SQAMGRVGEVITRTWQTADKNKKEFGRLKEEKDNDNFRIKRYLSKYTNPAIAHGISEY 422
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 422 IGSVEEGKIADLVVWNPAPFGVKPKIVIKGMVVFSEMGSDSNASVPTPQVYVREMFHGH 481
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 423 VGSVEVGKVLVLVWSPAFFGKPNMIKGGFIALSQMGDANASIPTPQVYVREMFHAH 482
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 482 GKAKEDTSTITFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 483 GKAKYDANITFVSQAAYDKGIEELGLERQVLVKNCRNITKKDMQFNDTTAHIEVNPET 542
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Qy 542 FEVFDGKLCSTKPTSOVPLAQRVTF 568
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 543 YHVFVDGKEVTSKPANKVSLAQLFSIF 569
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

RESULT 10

```
US-10-335-977-8611
; Sequence 8611, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:14:48 ; Search time 7.15365 Seconds
(without alignments)
240.659 Million cell updates/sec

Title: US-09-904-994B-3

Perfect score: 2999

Sequence: 1 MKKKQYVNTYPTKGDV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.New.*

- 1: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pap.*
- 2: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pap.*
- 3: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pap.*
- 4: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pap.*
- 5: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pap.*
- 6: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pap.*
- 7: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pap.*
- 8: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789.5	59.7	571	1	US-10-793-626-118
2	1380.5	46.0	431	1	US-10-485-517-182
3	123.5	4.1	802	1	US-10-510-386-2
4	113	3.8	2314	7	US-11-013-759-11
5	105.5	3.5	801	1	US-10-793-626-90
6	100	3.3	403	7	US-11-074-176-354
7	100	3.3	408	7	US-11-074-176-186
8	99.5	3.3	1992	7	US-11-013-759-3
9	99.5	3.3	1992	7	US-11-013-759-13
10	99.5	3.3	2047	7	US-11-013-759-4
11	99.5	3.3	2047	7	US-11-013-759-7
12	96	3.2	1160	1	US-10-131-826A-234
13	94.5	3.2	480	1	US-10-821-234-1465
14	94	3.1	2053	7	US-11-013-759-9
15	93	3.1	1047	1	US-10-510-386-200
16	93	3.1	1437	7	US-11-074-176-96
17	90.5	3.0	585	1	US-10-793-626-2124
18	90	3.0	547	7	US-11-082-389-182
19	89.5	3.0	491	1	US-10-793-626-2808
20	89	3.0	770	1	US-10-821-234-1269
21	88	2.9	2323	1	US-10-793-626-760
22	86	2.9	469	1	US-10-793-626-2204
23	85.5	2.9	1267	7	US-11-109-156-35
24	84.5	2.8	530	1	US-10-858-730-67
25	84.5	2.8	3717	1	US-10-821-234-1076

ALIGNMENTS

RESULT 1

US-10-793-626-118

; Sequence 118 Application US/10793626

; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 118

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: amino acid sequence

; OTHER INFORMATION:

US-10-793-626-118

Query Match

Best Local Similarity 59.7%; Score 1789.5; DB 1; Length 571;

Matches 332; Conservative 92; Mismatches 141; Indels 5; Gaps 2;

Qy

2

KMKKQYVNTYPTKGDVRLGDTDLMAEVEHDTTYGEEELKFGAGKTIREGMGQS--N 58

Db

4

KMTQSQTSLSYPTGPTGDSVRLGDTNLFQAQVEKDYANYGDEATFGGKKSIRDMGAPNVT 63

Qy

59

SPDENTLDLVITNAMIIDYTGIIYKADIGIKNGKHIGIKAGKNQMDGVSPHMYVGCTE 118

Db

64

RDDKNVADLVLTNALIIDYDKIVKADIGIKNGYIFKIGKAGNPIDMNDV--IIGATTD 121

Qy

119

ALAGEGMIITAGGDSHTFLSPQFPPTALANGVTTFMGGGTGPVDTGNATTITPGKNWL 178

Db

122

IIAAGKIVTAGGIDTTHVHFINPEQAEVALBSGITHHGGGTGASEGAKATITVTPGPHI 181

Qy

179

HRMLRAAEYSMNVGFLLGKGNSSKKQIIVEQVAGAIQFKLHEDWGTTFPSAIDHCLSVAD 238

Db

182

HRMLEAAEEMPINVGFTKGQAVNHTALIEIQHAGAIGLVKHEDWGTGATPSALSHALDVAD 241

Qy

239

EYDVQVCITDVTNVEAGYVDDTLNAMNGRAIHAYHIEGAGGGHSPDVTIMAGELNIPSS 298

Db

242

EFDVQVALHADTLEAGFMEDTMAAVKDRVLHMYHTEGAGGGHAPDLIKSAAYSNIIPSS 301

Qy

299

TTPTIPTINTVAEHLDMLTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAVT 358

Db 602 SYGTFL--KDKGKQTKQAFTIENLSSHRKAY---QLEYSFKGTGITVS---GTERVVVPA 654
QY 469 PQVVYVYREMGHGGKAKFDTSITFSKVAYENGVEKELGL-----ERQVLPVKN 518
Db 655 NQT-----GKAATV-----NSAKTKAGTYEGTVYIREDGRKVAEIPLL 695
QY 519 RNITKDKDFNDKTIKIVDP-----KTFEVFDG 548
Db 696 LIVKEPDY---PRVTSVTVBPGAKQAGAYTIEAYLPG 728

RESULT 4
US-11-013-759-11
; Sequence 11, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-92MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-11

Query Match 3.8%; Score 113; DB 7; Length 2314;
Best Local Similarity 20.0%; Pred. No. 0.47;
Matches 128; Conservative 68; Mismatches 201; Indels 244; Gaps 31;
QY 8 YNTY-----GPTGDKVRLGDTDLWAEVHDYTYGE--ELKFGAGKTIREGMCQSN 58
Db 779 FVSTYNTVDPIDGNATTAFTYDETNTQTSKVTYDNNVDEKTIETLTDGNGKTNIGV--- 834
QY 59 SPENTLDLIVITNAMIIDYGIYKADIGIKNGIKHIGKAGN-----KMDQGVSP--- 109
Db 835 -----KTTTLTTTNA-----NGKATNFSTTDNALVNAKDIAENLNTLAK 874
QY 110 --HMVVGVGTEAL-----AGEGMIITAGGIDSHTHPLSPQFPPTALANGVTTMFGG 158
Db 875 EIHHTKGTADTALQTFKVKDKGATDDETITVG-----KDGTONKTV--- 916
QY 159 GTPGVDTNATTTTPGK-----WNLHRLRAAEYSNMVGFGLKGNSSSKQLVEQVE 211
Db 917 NTLKLGKGLTVATNKDGTVTFTGINTQSLKAGDSTTLNKGDLGSLKNPASN-----BQIQ 972
QY 212 AGAIGFKLHE--DNGTTPSAIDHCLSVADEYDVQVCIHTDTVNEAGYVDDTLNAMNGRAIH 270
Db 973 VGADGVKFAKVDKGNSTGIDGTSRI-----TKDQIGFT----- 1006
QY 271 AVHIEGAGGSHSPDVTMAGELNLSPTTPTIPTINTVAEHLDMLTCHLDKRIRED 330
Db 1007 -----GANG-----SLDTTKPHLT-----DKLVGE 1028
QY 331 LQFSQSRIRPGSIAEDVLHDMGVIAWTSSDSQAMGRAGEV-----IPRTWQTADKNK 383
Db 1029 VEITNTGINAGKKITNI--QSGDITQNSNDVATGGRVYDLKTELESKINSAAKTAQNSL 1086
QY 384 KEFGKLPEDGKNDNFRPRIKYISKYITINPALTHGVSEYIGSVBEGKIADLVWNPAPFGV 443
Db 1087 HEFSDVADEQC---NHFTVSNPYSSYD-----TSKTSDVITFAGENGITTKVNGGVVRGI 1138
QY 444 -----KPKIVIKG---GMVVFSEMGDSN-----ASVPTPQPVYVREMGH---H 481

Db 1139 DOTKGLTTPKLTGVNNGKGIVIDSKDQONTITGLSNTLANVT-----NDGAGHALSQ 1191
QY 482 GKAKFDT-----SITFVS-----KVAYENGVEKELGLERQVLVPV 515
Db 1192 GLAN-DTDKTRAASIGDVLNAGFNLQNGNEAVDFVSTYDTVTDFIDG----- 1236
QY 516 KNCENITKDKPFND--KTAK-----ITVDPKTFEFVVDGKL 550
Db 1237 ----NATTAKVTYDDTSTKTSKVVDVNVNDNKTIETVSDKKL 1273

RESULT 5
US-10-793-626-90
; Sequence 90, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 90
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-90

Query Match 3.5%; Score 105.5; DB 1; Length 801;
Best Local Similarity 19.5%; Pred. No. 0.39;
Matches 128; Conservative 75; Mismatches 244; Indels 211; Gaps 29;
QY 22 LGDTDLWAEVHDYTYGEEELKFGAGKTIREGMCQSNPDENTLDLVTNAMIIDYTYG 81
Db 214 IGDKDV-----EALKLDGRVIE-----TDLVMAVGIRPYTEVA 248
QY 82 KADTGI-----KNGKHIGIGKAGNKD-----MQDGVSPHVMVVG 116
Db 249 K-DSGLDVNRGIVNDYMQTSDSHIYAVGECAEHDGKYGVLVAPLYEQGKVLADYLTGE 307
QY 117 TEALAGSGM-----IITAGI--DSHTH-----FLSQOQPPTALA 149
Db 308 TKYKGSSTTTSLKVSGLDYSAGQIVDEDDVHGVEIFNSVDNIYKYVLSQGVVGA 367
QY 150 NGVTTFMGGTGTPVDGNTNATITPKKNLHRLRAAEYSNMVGFGLKGNSSSKQLVEQ 209
Db 368 YGDTD-----DGRF-----YNNMKKHETLEDYTL-VSLHKGDEDEAGTSIADM 410
QY 210 VEAGAIGFKLHEDMGTTPSAI--DHCLSVADEYDVQVCIHTDTVNEAGYVDDTLNAMNGRA 268
Db 411 SDEFTICGCGVDKGTIVNAITSKGLTSVDE-----VTKATKAGNSCGKKGQIGELLO 466
QY 269 IHAYHIEGAGGGHSP-----DVTMAGELNLPSSTT-----PTIPY 305
Db 467 LGDDFIAAKPTGICPCTDLTRDQIVTQIRAKNLKSSKEVRHVLDFDKDKGCPKCPAINY 526
QY 306 TINTV--AEHLDMLTCHLDK-----IREDLQFSQSRIRPGSIAEDVLHDMGV 354
Db 527 YLNMVYFPEHRD-----EKDSRFANERYHANIQNDGTFSVIPQMRGGVTDADQJIRLGE 580
QY 355 IA-----WTSSDSQAMGRAG---EVIPTWQTAD-KNKEFGKLPEDGKNDNFRPRIK 403
Db 581 VAKYNVPLVKVTSQVRVGLYGLKKEELPQWKLGMRSASAYKKTYSKSCVGEKPCR 640
QY 404 YISKYITINPAL-----THGVSEYIGSVBEGKIADLVWNPAPFGVVKPKTV 448
Db 641 FGTCYTTLRGLRLEKTFEYIDTPHKFKMGVSGCPSPSCVSEGVKD-----FGV---IS 689

Qy	449	IKGKVVVF-----SEMGSNASVPTQPV-----YYRMFGHHGKAKEDTSTT	491
Db	690	VENGYQIFIGNGGTDVTVGKLLTTVETEDSVIQLCGALMQYYRET-GVYA-----ERTAP	744
Qy	492	FVSKVAYVENGKELGLGLERQVLPVKNCERNITKKQPFKNDKTAKITVDPRTFFVFDGK	549
Db	745	WLERMGFENVKVLNLNOBKQ-----KELYLRIMEAKKAVENPWTETVENK	790

RESULT 6

```

US-11-074-176-354
; Sequence 354, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 403
; TYPE: PR1
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-354

```

[illegible]

RESULT 7

```

US-11-074-176-186
; Sequence 186, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Ruetemann, Eric
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Peril, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694

```

```

; CURRENT APPLICATION NUMBER: US/11/074,176
;
; CURRENT FILING DATE: 2005-03-07
;
; PRIOR APPLICATION NUMBER: 60/551,161
;
; PRIOR FILING DATE: 2004-03-08
;
; NUMBER OF SEQ ID NOS: 381
;
; SOFTWARE: Fast-SEQ for Windows Version 4.0
;
; SEQ ID NO 186
;
; LENGTH: 408
;
; TYPE: PRT
;
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-186

```

Query Match	3.3%;	Score 100;	DB 7;	Length 408;
Best Local Similarity	24.6%;	Pred. No. 0.4;		
Matches	72;	Conservative 43;	Mismatches 112;	Indels 66; Gaps 18;
Qy	301	PTPYTTNTVAEHLDMMLTCHHLDKIREDLQFSQSRIRPGSIAAEVDVLHDMGVIA	TSS	360
Db	47	PTIKYII-----EHGKAILLSHLG-RVKSADAKKELSLKEVABRLSELL-DKPVTF	VP	100
Qy	361	DSQAMGRA-----GEVI-----PRTWOTADKN--KKSEFGKLPEDGK-----	DN	400
Db	101	EKGEVEEAIINMKDGVVVLENTRFQIDINDPKRSGNDPKLGEYASLGDFVND	AF	160
Qy	401	I--KRYISKYTTINPAL--THGVSVEYIGSEGGKIADL--VVMNPAFFGVKPKVIKGM	VM	455
Db	161	TAHRSHASNVGIATAMKENGKPAAGVYLKEIKYLGDAVDP-----VHFPVILG	AKV	216
Qy	456	FSEMGDSNASVPTPQPVY-----YREMGHH--GKAKEDTTSITFVSKVAYENG	VKEKL	506
Db	217	SDKIGVIENLIPKSDHILIGGMAYTFLAAQGHKIGKSLFE-----ADKVDL	AKELLEKA	271
Qy	507	GLERQVLVPKNCRNITIKDPKFNDDTKAKITVD-----PKTFE	VFVD	547
Db	272	G-DKILPVD---NVAATEFS-INDASREVVGDDIPDNMMGLDIPGKTI	IAFKD	319

RESULT 8

```

US-11-013-759-3
; Sequence 3, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-3

```

Query Match	3.3%	Score 99.5	DB 7	Length 1992
Best Local Similarity	21.8%	Pred. No. 4.4		
Matches 114	Conservative 49	Mismatches 160	Indels 201	Gaps 30

Qy	54	MQQNSPDE-NTLDLVTNAMI	DTGYIYKADIGIKNGKI-HGIG-----KAGNKMDQD	106
	:			
	:			
Db	921	VGQKXANNQVNTLT	TLKGENGLNI-----KTD--KNGTVTFGINTTSGLKAGKSTLNDG	971
	:			
	:			
Qy	107	VSPHNVGVGTEALAGSGMI-ITAGGIDSHTHFLSPQ	PFALANGVTTMFGGGTGPDVG	165
	:			
	:			
Db	972	-----GLSKNPTGSEOT	OVGADGV-----KPAKVNNNGVV-----CAG-IDG	1008
	:			
	:			

QY 166 TNATTITPKGNLHRLMLRAAEYSYMNVLGKGNSSKK-----QLVQVEA 212
Db 1009 T--TRITRDE-----IGFTGTNGSLDKSPHLSKDGINAGGKKTNIQS 1050
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
Db 1051 GEIAQNSHD--AVTGKGI-----YDLK-----TELENKISSAKT--AQN--SLHEF 1091
QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
Db 1092 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGDQTKGLTTPKL 1151
QY 318 MTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
Db 1152 TVGNNGKGVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1198
QY 378 TADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEYIGSVSEGGKIADLVVWN 437
Db 1199 -----KDEDKTR-----AASIVD----- 1211
QY 438 PAFFGVKPKIVIKGVMVFSEMGDSNA--SVPTPOPVYVREMFGHHGKAKF---DTSITF 492
Db 1212 -----VLSAG---FNLOGGEAVDFVSTYDVTNFPAD--GNATTAKVYDDTSKT- 1255
QY 493 VSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKNDKTAKIT 536
Db 1256 -SKVYDVNVDD-----TTIEVKDKKLGKVTITLT 1284
RESULT 9
US-11-013-759-13
; Sequence 13, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1992
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-13
Query Match 3.3%; Score 99.5; DB 7; Length 1992;
Best Local Similarity 21.8%; Pred. No. 4.4;
Matches 114; Conservative 49; Mismatches 160; Indels 201; Gaps 30;
QY 54 MGQSNPDE-NTLDLVIITNAMIIDYTGIVKADIGIKNGKI-HGIG-----KAGNKDMQDG 106
Db 921 VQKQNNQVNTLTGKGNLNI-----KTD---KNGVTFTGINTTSLGKAGKSTLNDG 971
QY 107 VSPHVMVGVGTTEALAGEGMI-ITAGGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPVDG 165
Db 972 -----GLSIKNPTGSEQIQVGADGV-----KFAKVNNGVV-----GAG-IDG 1008
QY 166 TNATTITPKGNLHRLMLRAAEYSYMNVLGKGNSSKK-----QLVQVEA 212
Db 1009 T--TRITRDE-----IGFTGTNGSLDKSPHLSKDGINAGGKKTNIQS 1050
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
Db 1051 GEIAQNSHD--AVTGKGI-----YDLK-----TELENKISSAKT--AQN--SLHEF 1091

QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
Db 1092 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGDQTKGLTTPKL 1151
QY 318 MTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
Db 1152 TVGNNGKGVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1198
QY 378 TADKNKKEFGKLPEDGKNDNFRIKRYISKYITNPALTHGVSEYIGSVSEGGKIADLVVWN 437
Db 1199 -----KDEDKTR-----AASIVD----- 1211
QY 438 PAFFGVKPKIVIKGVMVFSEMGDSNA--SVPTPOPVYVREMFGHHGKAKF---DTSITF 492
Db 1212 -----VLSAG---FNLOGGEAVDFVSTYDVTNFPAD--GNATTAKVYDDTSKT- 1255
QY 493 VSKVAYENGVEKGLGLERQVLPVKNCRNITKKDFKNDKTAKIT 536
Db 1256 -SKVYDVNVDD-----TTIEVKDKKLGKVTITLT 1284
RESULT 10
US-11-013-759-4
; Sequence 4, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-4
Query Match 3.3%; Score 99.5; DB 7; Length 2047;
Best Local Similarity 21.8%; Pred. No. 4.6;
Matches 114; Conservative 49; Mismatches 160; Indels 201; Gaps 30;
QY 54 MGQSNPDE-NTLDLVIITNAMIIDYTGIVKADIGIKNGKI-HGIG-----KAGNKDMQDG 106
Db 976 VQKQNNQVNTLTGKGNLNI-----KTD---KNGVTFTGINTTSLGKAGKSTLNDG 1026
QY 107 VSPHVMVGVGTTEALAGEGMI-ITAGGIDSHTHFLSPQQPPTALANGVTTMFGGCTGPVDG 165
Db 1027 -----GLSIKNPTGSEQIQVGADGV-----KFAKVNNGVV-----GAG-IDG 1063
QY 166 TNATTITPKGNLHRLMLRAAEYSYMNVLGKGNSSKK-----QLVQVEA 212
Db 1064 T--TRITRDE-----IGFTGTNGSLDKSPHLSKDGINAGGKKTNIQS 1105
QY 213 GAIGFKLHEDWGTTPSAIDHCLSLVADEYDVQVCIHTDTVNEAGYVDDTLNMGRAIHAY 272
Db 1106 GEIAQNSHD--AVTGKGI-----YDLK-----TELENKISSAKT--AQN--SLHEF 1146
QY 273 HIEGAGGGH-----SPDVIITMAGELNLPSSSTTPTIPTYIN--TVAEHLMDL 317
Db 1147 SVADQGNFTVSNPNYSSYDTSKTSVDITFAGENGITTKVKNKGVRVGDQTKGLTTPKL 1206
QY 318 MTCHHLDKRIREDLQFSQSRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGRAGEVIPRTWQ 377
Db 1207 TVGNNGKGVIDSQNGQNTITGLSNTLANVNDKGSVRTTEQ-----GNII----- 1253

QY 378 TADKNNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEYIGSVBEGKIADLVVWN 437
Db 1254 -----KQEDKTR-----AASIVD----- 1266
QY 438 PAFFGVKPKIVIKGWMVFSEMGDSNA--SVTPQPVVYVREMGHHGAKF---DTSITF 492
Db 1267 -----VLSAG---FNLOGNGEAVDFVSYDTYVNFAD--GNATTAKVTYDDTSKT- 1310
QY 493 VSKVAYENGVEKGLGLERQVLVKNCRNITKKDFKFNKTAKIT 536
Db 1311 -SKVVYDVNVDD-----TTIEVKDKKLGKVTTLT 1339
RESULT 11
US-11-013-759-7
; Sequence 7, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA
; FILE REFERENCE: 1038-921MIS:jb
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/11/013,759
; PRIOR FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-7
Query Match 3.3%; Score 99.5; DB 7; Length 2047;
Best Local Similarity 21.8%; Pred. No. 4.6;
Matches 114; Conservative 49; Mismatches 160; Indels 201; Gaps 30;
QY 54 MQQSNPDE-NTLDLVITNAMIIDYTGIIYKADIGIKNGKI-HGIG-----KAGNKMDDQ 106
Db 976 VQKQNNQVNTLTGKGNLMI-----KTD--KNGVTFTGINTTSLGKACKSLNDG 1026
QY 107 VSPHVVGVGTALAGEMI-ITAGGIDSHTHFLSPQQPFTALANGVTTFMFGGCTGPVDG 165
Db 1027 -----GLSIRKPTGSEQIQVGADGV-----KFAKNVNNGV-----GAG-IDG 1063
QY 166 TNATTITPGKWNHLRMAAEYSVMVFLGKGNSSKK-----OLVQVEA 212
Db 1064 T--TRITRE-----IGFTGNSLDKSKPHLSKGINAGGKKTNIQS 1105
QY 213 GAIGFKLHEDWGTTPSAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLNANGRAIHAY 272
Db 1106 GEIAQNSHD--AVTGCKI-----YDLK-----TELENKISSTAKT--AQN--SLHEF 1146
QY 273 HIEGAGGGH-----SPDVTWAGELNLPSTTPTIPTYIN-TVAEHLMDL 317
Db 1147 SVADEQGNFTVSNPYSSYDTSKTSVDITPAGENGITTKVKNKGVVRGIDQTRGLTTPKL 1206
QY 318 MTCHLHDKREIDLOFSQSRIRPGSIAEDVLHDMGVIAMTSSDSQAMGRAGEVIRPTWQ 377
Db 1207 TVGNNGKGLVIDSQNGQNTITGLSNLTANVTNDKGSVRTTEQ-----GNII----- 1253
QY 378 TADKNNKEFGKLPEDGKNDNFRIKRYISKYITINPALTHGVSEYIGSVBEGKIADLVVWN 437
Db 1254 -----KQEDKTR-----AASIVD----- 1266
QY 438 PAFFGVKPKIVIKGWMVFSEMGDSNA--SVTPQPVVYVREMGHHGAKF---DTSITF 492
Db 1267 -----VLSAG---FNLOGNGEAVDFVSYDTYVNFAD--GNATTAKVTYDDTSKT- 1310

QY 493 VSKVAYENGVEKGLGLERQVLVKNCRNITKKDFKFNKTAKIT 536
Db 1311 -SKVVYDVNVDD-----TTIEVKDKKLGKVTTLT 1339
RESULT 12
US-10-131-826A-234
; Sequence 234, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 234
; LENGTH: 1160
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-234
Query Match 3.2%; Score 96; DB 1; Length 1160;
Best Local Similarity 18.1%; Pred. No. 3.8;
Matches 99; Conservative 81; Mismatches 162; Indels 206; Gaps 31;
QY 16 KGDVRLGDT---DLMAVEHEDYTYGEEELKFGA-----GKTIREGMOQSNSPDNTLDL 67
Db 66 KSDKNRIGGTYYKTKIYKEYKDD--SYTDEVAQPAWLGFLGVLQAEVG-----DV 113
QY 68 VI-----TNAMIDYTG- YKADIGIKNGKIHIKAGKMDQGVSPHVVGVGTEA 119
Db 114 ILIHLNKFATRPYTIHPHFVFEKD---SEGSLYPDGSSGPLKADDSVPP----- 160
QY 120 LAGEGMIITAGGIDSHTH-FLSPQQFPTALANGVTTFMFGGCTGPVDGNTATTITPGKWNL 178

Db 161 -----GG--SHLYNWTIPE-----CHAPTDADPACL-----TWIY 189
QY 179 HRMLRAAEYSMMVGFILG-----KGNSSSKQ-----LVQVEAGAGIFKXHE 221
Db 189 HSHVDAPRDIA--TGLIGPLITCKRGALDGNSPQROVDHDFLLFSVVDENLSWHLNE 246
QY 222 DWGT---TPSAIDHCLSVADYDVQVCIHTDTVNEAGYVDDTLNAMN--GRAIHAYHIEG 276
Db 247 NIATYCSDPASVD-----KEDETFOESNRMHAIN--GFVFGNLPCLNMCAKRVAMHLFG 299
QY 277 AGG-----CHSPDVTMAGEINILPSS--TTPTIPTYTINTVAEHLDM 317
Db 300 MGNEIDVHTAFFHGMQLTTRGHHTDV-----ANIFPATTVAEMVWPEPGT-----WL 347
QY 318 MTCHLDKRIREDLQ-----FSQSRIRPGSIAAEDVLHDMGVIAMTSSDS 362
Db 348 ISC-QVNSHFRDGMQALYKVKSCSMAPPVLLTKVKRQYFIEAHEIQWDYGPWCHDSTG 406
QY 363 QAMGRAGEVIPRTWQTADKNKEFG-----KLPEDKGDNDFRIKRYISKYTIINPALTHGV 418
Db 407 KNLREPGSISDKFFQ---KSSSRIGGTYWKVRYEAFQDETFOEKHLEE-----452
QY 419 SEVIGSVBECKIADLVVNPAPFGVKPKIVIKGMMVVFSEMGD-----SNASVP---T 468
Db 453 DRHLGIL-----GPVIRAEVGDITQVVFYFNRSQPFPSMQ 486
QY 469 PQPVYVRE 476
Db 487 FHGVFYEK 494

RESULT 13
US-10-821-234-1465
; Sequence 1465, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 1465
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1465

Query Match 3.2%; Score 94.5; DB 1; Length 480;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 66; Conservative 48; Mismatches 95; Indels 127; Gaps 18;
QY 50 IREGMGQSNPDENTLDLVTNAMIIDYTGIVKADIGIKNGKHGIGKAGNKDMQD--- 105
Db 201 VASGSGFIVSED---GLIVTNAHV--TNKRVKVELKNGAY---EAKIKDVEKADI 251
QY 106 -----GVSHPMVGVGTEALAGEMIITAGIDSHTHFLSPQFPPTALANGVTT--- 154
Db 252 ALIKIDHOGKLPVLLGRSSELPGE-FVVAIGS-----PFSLQNTVTTGIV 297
QY 155 -----MFGSGTGP-----VDGTNATTITPG-----K 175
Db 298 STTQGGKELGRNSDMDYIQTDAIYNGSGGGLVNLGDEVIGINTLKVTAISFAIPS 357
QY 176 WNLHMLRAAEYSMMVGFILGKGNSSKKQLVQVEAGAGIFKXHE-----HE 221

Db 358 DKIKKFLTESHDRO-----AKGKAITKKY-----IGIRMSLTSSKAKELKDRHR 403
QY 222 DWGTPPSAIDHCLSVADYDVQVCIHTDTVNEAGYV--DDTLNAMNGRAIHAYHIEGAGG 279
Db 404 DF---PDVI-----SGAYIIEVI--PDTPAEAGGLKENDVIIISINGQSV-----VSA 445
QY 280 GHSPDVTMAGEINILPSSTTPTIPTYTINTVAEHL 315
Db 446 NDVSDVIKRESLNNVVRGNEIDMITV--IPEEID 479

RESULT 14
US-11-013-759-9
; Sequence 9, Application US/11013759
; Publication No. US20050249747A1
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Yan Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN OF MORAXELLA
; FILE REFERENCE: 1038-921MIS-1b
; CURRENT APPLICATION NUMBER: US/11/013,759
; CURRENT FILING DATE: 2004-12-16
; PRIOR FILING DATE: US/09/361,619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-759-9

Query Match 3.1%; Score 94; DB 7; Length 2053;
Best Local Similarity 19.0%; Pred. No. 13;
Matches 120; Conservative 85; Mismatches 210; Indels 218; Gaps 31;
QY 13 GPTKGDVKRIGDITDLMAEVEHDYTY-----GELKFGAGKTIR-EGMGQSNPDENT 64
Db 341 GSRDITDANVAQLKLVLELANRKITFKGDNNSVERGLNTLTIKGDAOTNALTEAN 400
QY 65 LDLV-----ITNAMIIDYGIYKADI-----GIKNGK- 91
Db 401 IGVTDGNGLKVKLAKELTGLTSVSATNKITVSNTNNNAELQSGGLTFSPITGYTKDT 460
QY 92 ---IHG-----IGKAGNKDMQDGVSPHM---VVGVTBALAGE 123
Db 461 VYSIDGLKFTNDSNSIATKGTTRITKKGIFAGTNDGVDESKPYLDNEKLVGNSTLNSG 520
QY 124 GMII--TAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATITPGKWNLHRM 181
Db 521 SLTVNNTTGN-----KQIQVG-ANGI--KEATVANNVANTSATVGT-----558
QY 182 LRAAEYSMMVGFILGKGNSSKKQLV---EOVEAGAIGFKLHEDMGTTPSAIDHCLSVAD 238
Db 559 ARITEE---KIGFAGTNDGVDEQAPYLDKERLKVGRVEI-----TTDSGIN-----AG 603
QY 239 EYDV---QVCIH7D--TVNEAGYVDDTLNAMNGRAIHAYHIEGAGGHSPPDVTMAGEL 292
Db 604 NHKITGLTNGTANTDAVTIKQDKAKPTLNAGDGISINS-----NNGDLVDSGNI 654
QY 293 NILPSSTTPTIPTYTINTV-----AEHLDMLTCHLDKRIREDLQFSQ 335
Db 655 -----TTPTYNISVKTTKLSNNGTSGNNKFSVSNADHNNSLVTAOKDLADYLNKVNETAD 708
QY 336 SRIRPGSIAAEDVLHDMGVIAMTSSDSQAMGAGEVIRPTWQTADKNKEFKLPEDKGD 395
Db 709 SALPSPFKVQNGD-----NSNNAITVGK-----DTNGKTFNTLKLKGN 746
QY 396 NDNPRIKRYISKYTIINPALTHGVSE---YIGSVBECKIADLVVNM--PAFFGVKPKIVIKG 451

Db 747 GVNITNRTATGTVTFCIDOSNGLTTPKLTGVSNTGN--RLVIEQVPSADGNSTKNIIG 804
QY 452 -----GMVFSEMGDSNASVPTQPVPVYREMFHGHGKAK-----FDT 488
Db 805 LSPFLPSIASPSGRNIALGNTIIEKDKSNAASIDDLNAGFN--LKNNGKDKDFVSTYDT 862
QY 489 -----SITP-----VSKVAYENGVEK 505
Db 863 VDFIDGNATTATVTDYDEANQTSKVAYDVNVDEK 895

RESULT 15

US-10-510-386-200
; Sequence 200, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 200
; LENGTH: 1047
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-200

Query Match 3.1%; Score 93; DB 1; Length 1047;

Best Local Similarity 19.2%; Pred. No. 5.7; Indels 156; Gaps 27;
Matches 109; Conservative 75; Mismatches 227

QY 42 LKFGAGKTI-----REGMGQNSP--DENTLDLVITN--AMIIDYTIYKA-DIGIKN 89
Db 562 LKVAASKNVTIIVATGNDGVSEISYPASSKYTLISVGATNNLDIVSDYSNYGKGLDMVAPG 621
QY 90 GKTHGIGKAGKMDQDQGVs---PHMVVGTEALAGEMIITAGGIDSHTHFLSPQOFPT 146
Db 622 TDIPSLVPDGNVTYMGTSMAAPHVA-----AAAGLLL-----SQNP SLKPQIAS 667
QY 147 ALANGVTTMFGGTPGPDGNTATITPGKNLHRLRAAEYSNMGVFLCKGNSSSKQL 206
Db 668 LLT-----ETTADVAFEEQDNPNDYDL 690
QY 207 VEQVEAGAIGFLHEDWGTTPSAIDHCLSVADYDVQVCIH----- 247
Db 691 DIEPAAQIPGYDFVSGWGL--NVFHAASV---FELNMKVHPVLNRHTAVTGTAKSGVTV 745
QY 248 -----TDTVNEAGYVDDTLNAMN--GRAIHAYHIEGAGGCHSPD-----VITMAGE 291
Db 746 KILRGQVLGTGTAGKSGAFSVKIPAKQAGQVLHV-----AASGHQAETSRLTVVEKAPK 800
QY 292 LNLPSSTTPTIPTYTINTVAEH--LDMLMTCCHLDKRIREDLQFS-----QSRIRPG 341
Db 801 NPSVKRITNKDТАVGTGRTAAGYTIKVNACKKVIAQGRADASVSVKVINKQKEYAVLYV 860
QY 342 SIAAEDVLHDWGIAMTSSDSQAMGRAGEVI PRTWOTADKNKKEFGKLPEDGKDNDFRI 401
Db 861 SASADD--HRESGDVKMTADVIFPG-----APKVQVSDKS-----TVIQGTEANAQV 908
QY 402 KRYISKYTNIPALTHGVSEYIGSVEGKIADLVVMNPAFFGVKPKIVIKGMV-----VFS 457
Db 909 SAKAKGTIASGKANGKEYKLSRQKAGTVI-----GVTAKE--DKAGNVSKATAVT 959
QY 458 EMGDSNASVPTQPVPVYREMFHGHGKAKFTDSTFVS--KVAYENGVEKLG-----LERQ 511
Db 960 VLDKTPPSAPKVPNTVKST-AVKGKAANAALIVKSGKKTIGTGKADKKGAPFVKIKQ 1018

QY 512 VLPVKNCRNITKKDFKEN-DKTAKITV 537
Db 1019 --KANTVLAVTAKDKAGNTSKVSKIY 1043
Search completed: November 28, 2005, 08:25:59
Job time : 9.15365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2005, 08:13:38 ; Search time 46.4987 Seconds
(without alignments)
1009.915 Million cell updates/sec

Title: US-09-904-994B-3
Perfect score: 2999
Sequence: 1 MKKKQEVYNTYPTKGDKV.....KLCTSKPTSQVPLAQRYTFF 568

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep:*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep:*
3: /cgn2_6/prodata/1/iaa/H COMB.pep:*
4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep:*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298.5	76.6	568	2 US-09-338-920B-12	Sequence 12, Appl
2	2297	76.6	569	2 US-09-338-920B-6	Sequence 6, Appli
3	2294	76.5	569	1 US-08-467-822-26	Sequence 26, Appl
4	2294	76.5	569	2 US-08-432-697-26	Sequence 26, Appl
5	2294	76.5	569	2 US-08-466-248-26	Sequence 26, Appl
6	2290	76.4	566	1 US-08-920-095-3	Sequence 3, Appli
7	2290	76.4	566	4 PCT-US96-05800-3	Sequence 3, Appli
8	2281	76.1	570	2 US-09-431-705-5	Sequence 5, Appli
9	2244	74.8	569	1 US-08-467-822-21	Sequence 21, Appl
10	2244	74.8	569	2 US-08-432-697-21	Sequence 21, Appl
11	2244	74.8	569	2 US-08-466-248-21	Sequence 21, Appl
12	2244	74.8	569	2 US-09-338-920B-10	Sequence 10, Appl
13	2233	74.5	559	2 US-09-338-920B-8	Sequence 8, Appli
14	2028.5	67.6	568	1 US-07-732-242C-3	Sequence 3, Appli
15	1828	61.0	566	2 US-09-252-991A-26887	Sequence 26887, A
16	1821	60.7	625	2 US-09-489-039A-9216	Sequence 9216, Ap
17	1818.5	60.6	569	2 US-09-543-681A-6029	Sequence 6029, Ap
18	1818	60.6	567	2 US-09-328-352-5912	Sequence 5912, Ap
19	1817	60.6	567	1 US-08-967-513-5	Sequence 5, Appli
20	1817	60.6	567	1 US-08-687-645B-5	Sequence 5, Appli
21	1813	60.5	569	1 US-08-467-822-27	Sequence 27, Appl
22	1813	60.5	569	2 US-08-432-697-27	Sequence 27, Appl
23	1813	60.5	569	2 US-08-466-248-27	Sequence 27, Appl
24	1789.5	59.7	571	2 US-09-710-279-118	Sequence 118, App
25	1788.5	59.6	573	2 US-09-134-001C-5026	Sequence 5026, Ap
26	1787.5	59.6	840	1 US-08-467-822-25	Sequence 25, Appl
27	1787.5	59.6	840	2 US-08-432-697-25	Sequence 25, Appl

28	1787.5	59.6	840	2 US-08-466-248-25	Sequence 25, Appl
29	1708	57.0	570	2 US-09-602-777A-14	Sequence 14, Appl
30	1619	54.0	534	2 US-09-602-777A-16	Sequence 16, Appl
31	1166	38.9	308	2 US-09-338-920B-14	Sequence 14, Appl
32	469	15.6	121	2 US-10-012-819-160	Sequence 160, App
33	189.5	6.3	457	2 US-08-415-658-21	Sequence 21, Appl
34	175.5	5.9	622	2 US-09-949-016-10606	Sequence 10606, A
35	173.5	5.8	572	2 US-09-949-016-6070	Sequence 6070, Ap
36	173.5	5.8	578	2 US-09-949-016-11503	Sequence 11503, A
37	171.5	5.7	572	2 US-09-702-705-1815	Sequence 1815, Ap
38	171.5	5.7	572	2 US-09-736-457-1815	Sequence 1815, Ap
39	171.5	5.7	572	2 US-09-671-325-1815	Sequence 1815, Ap
40	171.5	5.7	572	2 US-10-017-754-1815	Sequence 1815, Ap
41	160.5	5.4	460	1 US-08-289-709-1	Sequence 1, Appli
42	160.5	5.4	460	1 US-08-602-656-1	Sequence 1, Appli
43	160	5.3	507	2 US-09-252-991A-23541	Sequence 23541, A
44	159.5	5.3	593	2 US-09-949-016-10192	Sequence 10192, A
45	159	5.3	484	2 US-10-009-782A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-338-920B-12
; Sequence 12, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338,920B
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Helicobacter heilmannii
US-09-338-920B-12

Query Match 76.6%; Score 2298.5; DB 2; Length 568;
Best Local Similarity 73.7%; Pred. No. 7.7e-214;
Matches 418; Conservative 74; Mismatches 74; Indels 1; Gaps 1;

QY	2	KMKKQEVYNTYPTKGDKVRLGDTDLMAEVHDYTYGEEELKFGAGKTIREGMGQNSPD	61
Db	3	KISRKEVSMYGPPTGDKVRLGDTDLILEVHDCTTYGEEIKFGGKTIKRDGMGTNSPS	62
QY	62	ENTLDLVTNAMIIDYTGIVKADIGIKNGKIHGKAGNKMQDGVSPHVVGVGTEALA	121
Db	63	SHELDLVTNALIVDYTGIVKADIGIKNGKIHGKAGNKLQDQVCNRLCVGPATEALA	122
QY	122	GEGMIITAGDISHTFLSPQOFTALANGVTTFGGGTGPDGTNATITPGKNLHRM	181
Db	123	AEGLIIVTAGGIDTHIHFIISQQIPTAFASGITTTMGGTGPDGTNATITPGRNLKEM	182
QY	182	LRAAEEYSMVVGLFGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPPSAIDHCLSVADEYD	241
Db	183	LRASEYAMLVGLGKGNVSFPALIDOLEAGAIGFKLHEDWGSTPSPAINHALNTADKYD	242
QY	242	VQVCIHDTVNEAGYVDDTLNANNGRAIHAYHIEGAGGCHSPDVTIMAGELNILFSSSTPP	301
Db	243	VQVAIHDTLINEACGVEDTLEAIAIRTIHTHTEGAGGCHAPDVIKMAGEFNILPASTNP	302
QY	302	TIPYTINTVAEHLDMLTCHHLDKRIREDLOFQSRIRPGSIAAEDVLHDMGVIAWTSSD	361
Db	303	TIPPTKNTAEAHMDMLM-CHHLDKNKIEDVEFADSRIRPQTIAARDKLHDMGIFITSDD	361

```

362 QY 362 SOAMGRAGEVJ PRTWQTADKNKKKGFGL PEDGKONDNDNFR I KRYISKYITNPAL THGVSEY 421
362 Db 362 SOAMGRVGEVJ TRTWQTADKNKKKGFGLR PEKGDNDNFR I KRYISKYITNPAL THGISEY 421
362 QY 422 IGSVEEGKIADLVVNMPAFFGVPKPIV I KGGWVVFSEMGDSNASVPTQP VTYREMFGGH 481
362 Db 422 VGSVEVEGYADLVLWSPAFFGKPNW I I KGGFIALSQMGDANAS I PTPQVTYREMFGGH 481
362 QY 482 GKAKFDTSITFVSKVAYENGVEKGLGLERQVL PVKNCNRI TKKDFKNDKTAKITVDPKT 541
362 Db 482 GKAKFDTNITFVSQVAYENGIKHELGLQRVVL PVKNCNRI TKKDLKFNDVTAHIEVNPET 541
362 QY 542 FEVFDVGDKCTSKPTSQVPLAQRITFF 568
362 Db 542 YKKVVDGNEVTSHAADKLSLAQLYNLF 568

RESULT 2
US-09-338-920B-6
; Sequence 6, Application US/09338920B
; Patent No. 6709851
; GENERAL INFORMATION:
; APPLICANT: Soman, Gopalan
; APPLICANT: Thomas, William D., Jr.
; APPLICANT: Monath, Thomas P.
; TITLE OF INVENTION: Stabilization of Helicobacter Urease
; FILE REFERENCE: 06132/023002
; CURRENT APPLICATION NUMBER: US/09/338, 920B
; CURRENT FILING DATE: 2002-12-23
; PRIORITY APPLICATION NUMBER: US 08/928,081
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-09-338-920B-6

```

```

Db      423   VGSVEVGKADLVLWSAPFAFGVKNMIIKGGFIALSQMGDANASITPTDPQVVYREMFHH    482
Qy      482   GKAKFDTSITFSVKAVAYENGVKELGLERQVLPVXNCRNITKKDFKFNDKTAKITVDPKT    541
          |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      483   GKAKYDANIIFVSQAAYDKGIGKEELGLERQVLPVKNCRNITKKWQFNDDTTAHIEVNPEP    542
          |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Qy      542   FEVPVDGKLCTSKPTSQVPLAQRITYFF    568
          :|||||:|||::|||::|||::|||::|
Db      543   YHVFVDGKVETSKPANVSLAQFLSIF    569
          :|||||:|||::|||::|||::|||::|

RESULT 3
US-08-467-822-26
; Sequence 26, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/POCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-467-822-26

```


QY	242	VQVCIHDTTVNEAGYVDDTLNAMNGRAHYHIEGAGGGHSDDVITWAGELNILPSSTTP	300
DB	243	VQVAIHDTTLNEAGCVDTTAAIAGRTWHTFTEGAGGGHAPDIIKVAGEHNILPASTNP	302
QY	302	TIPVTINTVAEHLDMLTCHLHKRIREDLOFSQSRIRPGSTAABDVLHDMGVIAWMTSSD	361
DB	303	TIPPTVTEAHHMDLMMVCHLHDKSIKSDVQPADSRIRPQTIAABEDTLHDMGIFSISSD	362
QY	362	SOAMGRAGEVIPRTWQTADNKKKEFGKLPEDGKONDNPRIKRYISKYTTINPALTHGVSEY	421
DB	363	SOAMGRVGEVITRTWQTADNKKKEFGRLKEBKGDNDNPFRIKRYLSKYTTINPAIAHGISEY	422
QY	422	IGSVEBGIADLVVWNPAPFGVKPIVITKGMVWFSEMGDSNASVPTPOPVYVREMPFGHH	481
DB	423	VGSVEVGKADLVLSWPAFFGVKPNMIKGGFIALSQMGDANASIPTPOPVYVREMPFAHH	482
QY	482	GKAKFDSTISFVSKVAVYENGVEKLGRLGRQVLVKNCRNITKKDKFENDTKAKITVDPKT	541
DB	483	GKAKYDANITFVSQAADYKGIKEBGLGRQVLVKNCRNITKKDMQFNDTTAHIEVNPET	542
QY	542	FEYFVDGKLGCTSKPTSQVPLAQRY	565
DB	543	YHVFVDGCKEYTSKPANKVSLAQLF	566
RESULT 7			
PCT-US96-05800-3			
; Sequence 3, Application PC/ITUS9605800			
; GENERAL INFORMATION:			
; APPLICANT: OraVax, Inc.			
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE			
; NUMBER OF SEQUENCES: 7			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Fish & Richardson P.C.			
; STREET: 225 Franklin Street			
; CITY: Boston			
; STATE: MA			
; COUNTRY: USA			
; ZIP: 02110-2804			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: PCT/US96/05800			
; FILING DATE: 23-APR-1996			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/431,041			
; FILING DATE: 28-APR-1995			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/568,122			
; FILING DATE: 06-DEC-1995			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Clark, Paul T.			
; REGISTRATION NUMBER: 30,162			
; REFERENCE/DOCKET NUMBER: 06132/020001			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (617) 542-5070			
; TELEFAX: (617) 542-8906			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 566 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
PCT-US96-05800-3			

Query Match

Best Local Similarity

76.4%; Score 2290; DB 4; Length 566;

73.6%; Pred. No. 5.1e-213;

Matches	415;	Conservative	75;	Mismatches	74;	Indels	0;	Gaps	0
Qy	2	KMKQEVVNTY	GPTGDKVRLGDTDLWAEV	EHDTYTYGEE	LKFAGKTI	IREGMQ	SQNSPD	61	
Db	3	KISKEVSMY	GPTTGDVKRLGDTDLIAE	VEHDYTYGEE	LKFGGKTI	IREGMQ	SQNSNPS	62	
Qy	62	ENTLDDLVT	NAMIIDYTYGIKADIGIKNG	KIHGIGKAGN	KMDQGVSPH	VMVVG	VTALA	121	
Db	63	KEEELDLIT	NALIVDYTYGIKADIGIKD	GKIAGIKGKN	KMDQGVKN	LSVGP	ATALA	122	
Qy	122	GEGMIITAG	GISDHTFLSPQFP	TALANGVTTMFG	CGTGVDC	TNATITIP	PGKWN	LHRM	181
Db	123	GEGLIIVAG	GIDTHIHIFISQO	IPTASGVTTM	GGTGPAD	GTNATITIP	GRNL	KWM	182
Qy	182	LRAAEEYS	MNVFLGKGNSSKKQ	LVEQVBAIGF	KLHEDWG	TTPSAID	HCL	SVAD	241
Db	183	LRAAEEYS	MNLGFLAKGNASND	ASLADQIEA	GAIKF	KIHEDWG	TTPS	AINH	242
Qy	242	VQVCIHDT	VNEAGYVDDTLN	MNGRAIHAYH	IEGAGG	HS	PDVIT	WAGEL	301
Db	243	VQVAIHDT	LTNEAGCVSDT	WAAIAGRT	MTHTFHT	EGAGG	HAPDI	IKVAG	302
Qy	302	TIPYTI	INTVAEHL	DMLTCHHL	DKIRIED	LQFS	RSIRPGS	IAAED	361
Db	303	TIPYTI	VNEAEH	DMLVCHHL	DKSIKED	VQFAD	SRIRPQ	TIAAED	362
Qy	362	SQAMGRAGE	IVPRTWOTADN	KNKEFGKL	PDGKDN	DNDFRI	KRYISKY	TIN	421
Db	363	SQAMGRGE	VITRTWOTADN	KNKEFGRL	KEEGK	DNDFRI	KRYLSKY	TIN	422
Qy	422	IGSVEE	GKIADLVVN	PAFQV	KPKIVIKG	MVVFSE	MGSD	SNASV	481
Db	423	VGSVEE	GKVADLVLS	PAFFQV	KPNMI	IKGFI	ALSQ	MGD	482
Qy	482	GKAKFDT	SITFVS	KVA	YENG	VKSKL	GLER	OVLP	541
Db	483	GKAKYANI	ITFVS	QAA	YDKG	IKGEL	GLER	QVLP	542
Qy	542	FEVEVDG	KLCTSK	PKTSQ	VP	LAQRY	565		
Db	543	YHVEVDG	KEVTSK	PANKV	SLA	QLF	566		
RESULT 8									
US-09-431-705-5									
; Sequence 5, Application US/09431705									
; Patent No. 6585975									
; GENERAL INFORMATION:									
; APPLICANT: Kleanthous, Harold									
; APPLICANT: Londono-Arcila, Patricia									
; APPLICANT: Freeman, Donna									
; TITLE OF INVENTION: Use of salmonella vectors for									
; TITLE OF INVENTION: vaccination against helicobacter infection									
; FILE REFERENCE: 06132/060001									
; CURRENT APPLICATION NUMBER: US/09/431,705									
; CURRENT FILING DATE: 1999-11-01									
; NUMBER OF SEQ ID NOS: 52									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 5									
; LENGTH: 570									
; TYPE: PRT									
; ORGANISM: Helicobacter pylori									
US-09-431-705-5									
Query Match									
Best Local Similarity									

Db 64 KEEELIITNALIVDTYGLYKADIGIKGKIAGIKGGKMDQGVKNLNSVGPATEALA 123
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATITPGKWNLHRM 181
Db 124 GEGLIIVTAGGIDTHIFHSFPQIPTAFASGVTTMIGGTGPADGTNATITPGRENKWM 183
Qy 182 LRAAEYSMNVLGFKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHDCLSVADRYD 241
Db 184 LRAAEYSMNVLGFLAKGNASNDASLADQTEAGAIGFAIHEDMGTPSPAINHALDVADKYD 243
Qy 242 VOVCHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 244 VQVAIATDTLNSAGCVEDTMAAIGRTMHTFTFTEGAGGSHAPDIIKVAHEHNILPASTNP 303
Qy 302 TIPTVINTVAEHLDMTCHLDKREIDLPQSQRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 304 TIPTVINTVAEHLDMVCHLDKREIDLPQADSRIRPQTAAEDTLHDMGIFSISSD 363
Qy 362 SOAMGRAGEVIRPTWOTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 364 SOAMGRAGEVIRPTWOTADKNKKEFKLKEEGKNDNFRIKRYISKYTINPALTHGVSEY 423
Qy 422 IGSVEEGKIADLVNPNPAFFGVKPKIVIKGGMVFSSEMCDNSASVPTPQPVYVREMFGHH 481
Db 424 VGSVEVGKADLVNWSAPFFGVKPNMIIKGGFIALSQMGDANASIPTPQPVYVREMFAHH 483
Qy 482 GKAKPDTSTTFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 484 GKAKYDANITTFVSOQAAYDKGIKEEGLERQVLVKNCRNITKDMQFNDTTAHIEVNPET 543
Qy 542 FEVFDGKLCTSKPTSOVPLAQRVTF 568
Db 544 YHVFVDGKEVTSKPNKVSQAQLFSIF 570

RESULT 9

US-08-467-822-21
; Sequence 21, Application US/08467822
; Patent No. 5843460
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferrero, Richard L.
; APPLICANT: Thiberge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSES: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/467,822
; APPLICATION NUMBER: US/08/467,822
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697

; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 569 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..569
; OTHER INFORMATION: /note= "URE B - FIGURE 3."
US-08-467-822-21
Query Match 74.8%; Score 2244; DB 1; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.5e-208; Mismatches 85; Indels 0; Gaps 0;
Matches 411; Conservative 71;
Qy 2 KKKQBYVNTYGTGDKVRLGDTDLMAEVEHDYTTYGEELKFGAGKTIREGMQSNPSD 61
Db 3 KISKEVSVSYGPTTGDVRLGDTDLILEVHEDCTTYGEEIKFGGKTIRDMGMSQTNPS 62
Qy 62 ENTLDLVITNAMIIDVTGIYKADIGIKNGIKHIGKAGNKMDDQGVSHVVGVTETALA 121
Db 63 SYELDLVITNALIVDTYGIYKADIGIKNGIKHIGKAGNKMDDQGVSHVVGVTETALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQFPPTALANGVTTMFGGTGPVDGTNATITPGKWNLHRM 181
Db 123 AGLIVTAGGIDTHIFHSFPQIPTAFASGVTTMIGGTGPADGTNATITPGRENKWM 182
Qy 182 LRAAEYSMNVLGFKGNSSSKQLVQVEAGAIGFKLHEDMGTPSPAIHDCLSVADRYD 241
Db 183 LRAAEYSMNVLGFLAKGNVSYEPSLRDQIEAGAIGFKLHEDMGTPPAIHLCLNVADRYD 242
Qy 242 VQVCIHTDTVNEAGVDDTLNAMNGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSSTP 301
Db 243 VQVAIHTDTLNEAGCVEDTLEAIGRTIHTFTFTEGAGGSHAPDVIKMAEFNLPASTNP 302
Qy 302 TIPTVINTVAEHLDMTCHLDKREIDLPQSQRIRPGSIAAEDVLHDMGVIAMTSSD 361
Db 303 TIPTVINTVAEHLDMVCHLDKREIDLPQADSRIRPQTAAEDTLHDMGIFSISSD 362
Qy 362 SOAMGRAGEVIRPTWOTADKNKKEFKLPEDGKNDNFRIKRYISKYTINPALTHGVSEY 421
Db 363 SOAMGRAGEVIRPTWOTADKNKKEFKLKEEGKNDNFRIKRYISKYTINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVNPNPAFFGVKPKIVIKGGMVFSSEMCDNSASVPTPQPVYVREMFGHH 481
Db 423 VGSVEVGKADLVNWSAPFFGVKPNMIIKGGFIALSQMGDANASIPTPQPVYVREMFGHH 482
Qy 482 GKAKPDTSTTFVSKVAYENGVEKGLERQVLVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKYDANITTFVSOQAAYDKGIKEEGLERQVLVKNCRNITKDMQFNDTTAHIEVNPET 542
Qy 542 FEVFDGKLCTSKPTSOVPLAQRVTF 568
Db 544 YHVFVDGKEVTSKAADELSQLYLNLF 569
RESULT 10
US-08-432-697-21
; Sequence 21, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes

APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 569 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..569
OTHER INFORMATION: /note="URE B - FIGURE 3."
US-08-432-697-21

Query Match 74.8%; Score 2244; DB 2; Length 569;
Best Local Similarity 72.5%; Pred. No. 1.5e-208;
Matches 411; Conservative 71; Mismatches 85; Indels 0; Gaps 0;
QY 2 KMKQBYVNTYPTKDKVRLGDTDLWAEVHDYTTYGEELKEGAGKTIREGQNSPD 61
DB 3 KISRKEYSMYGGTTGDRVRLGDTDLILEVHDCTTYGEEIKFGGKTIKRDGMSQNSPS 62
QY 62 ENTLDLVITNMTIDYTYGKADIGIKNGKIHGIGKAGNKDMQDGYSPHMVVGVEAL 121
DB 63 SYELDLVLTALIVDTYGYKADIGIKGKIGKAGNKDMQDGYDNNLCVGPATEAL 122
QY 122 GEGMIITAGGIDSHTFSPQPPTALANGVTTMFGGGTGPDVTNATTITPGKWLHRM 181
DB 123 AEGLIIVTAGGIDTHIFSPQIPTAFASGVTMTMGGTGPDGVTNATTITPGKWLKSM 182
QY 182 LRAEYYSNMVPLGKNGSSKQLVEQVEAGAIGFKLHEDWGTTTSPADHCLSVADEYD 241
DB 183 LRAAEYANMLGFLAKGNVSYPELSRDLQIEBAGAFKFIHEDWGTTTSPAAIHCLNVADEYD 242
QY 242 VQVCHTDTVNEAGYVDDTLNMMNGRAIHAYHIEGAGGHSDDVITMAGELNLTSPSTTP 301
DB 243 VQVAHTDTLNEAGCVEDTLEATAGTITHTFTTEGAGGGHAPDVIKWAGEFNLTLPASTNP 302
QY 302 TTPYTTNTVAEHLMDLMTCHHLDKRIREDLQFSQSRIPRPGSIAAEDVLHDMGVIAMTSSD 361
DB 303 TTPFTKNTAEHMDLMLVCHHLDSIKEDVQFADSRIRPQTIAAEDQLHDMGIFSITSSD 362

QY 362 SQAMGRAGEVTPRTWQTADKNKKEFGKLPEDGKDNDFRIKRIYSKYTINPALTHGVSEY 421
DB 363 SQAMGRVGYEITRTWQTADKNKKEFGRLKEKGNDNFRIRKRIYSKYTINPGIAHGISDY 422
QY 422 IGSVEEKGIALDLVWNPAFFGVKPKIVIKGGMVVFSEMGDSNASVPTPQPVYVREMFGHH 481
DB 423 VGSVEVGKYADLVLSWPAFFGKIKPMIIGGFIALSQMGDANASITPTQPVYVREMFGHH 482
QY 482 GKAKFDTSTITFVSKVAYENGVKELGLERQVLVKNCRNITKPKDFNDKTAKITVDPKT 541
DB 483 GKKNFDTNITFVSOAYKAGIKELGLDRAAPPVKNCNCRNITKDLKFNDVTAHIDVNPET 542
QY 542 FEVFDGKCTSKPTSQVPLAORYTFF 568
DB 543 YKVKVDGKEVTSKAADELSLAQLNLF 569
RESULT 11
US-08-466-248-21
Sequence 21, Application US/08466248
Patent No. 6258359
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastien
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
TITLE OF INVENTION: POLYPEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,248
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-02000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 569 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..569

Db 63 KEELDLIITNALIVDTYGIYKADIGIKGAGIKGGNKQTDQGVKNLNSVGPATEALA 122
Qy 122 GEGMIITAGGIDSHTHFLSPQOPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Db 123 GEGLIIVTAGGIDTHHFIHQIPAFASGVTTMIGGGTGPADGNTNATTITPGRNLRKM 182
Qy 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Db 183 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 242
Qy 242 VQVCHTDTVNEAGYVDDTLNANWGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSITP 301
Db 243 VQVAIHDTDLNAGCVEDTMAAIAGRTHMTYHTEGAGGGHAPDIIVKAGEHNILPASTNP 302
Qy 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIIRPGSAEEDVLHDMGVIAWTSDD 361
Db 303 TIPTVTNTEAEHMDMLMVCHELDKSIKEDVQFADSRIRPQTAAEDTLHDMGIFSTISSD 362
Qy 362 SQAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 363 SQAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 422
Qy 422 IGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 481
Db 423 VGSVEVGKVLADLVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 482
Qy 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 483 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 542
Qy 542 FEVFDGKLCIT 552
Db 543 YHVFVDGKEVT 553

RESULT 14

US-07-732-242C-3
; Sequence 3, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeashi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frishauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 568 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-732-242C-3

Query Match 67.6%; Score 2028.5; DB 1; Length 568;
Best Local Similarity 66.1%; Pred. No. 1.2e-187;
Matches 375; Conservative 77; Mismatches 112; Indels 3; Gaps 2;

Qy 3 MKQOEYVNTVPTKGDVKRIGDLDLWAEVHEHYTYTYGELKFGAGKTITREGMGQN-SPD 61
Db 4 MSRKQYADMFPTVGDAIRLADSELFTIEBKDYTYTYGDEVKFGGKIVRDGNGQPLATS 63
Qy 62 ENTLDLVTNAMIIDVTGIYKADIGIKGAGKAGKNDQGVSPHVMVGVTEALA 121
Db 64 DECVDLVTNAMIIDVTGIYKADIGIKGAGKAGKNDQGVSPHVMVGVTEALA 121
Qy 122 GEGMIITAGGIDSHTHFLSPQOPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Db 122 GEGMIITAGGIDSHTHFLSPQOPPTALANGVTTMFGGGTGPVDGNTNATTITPGKWNLRHM 181
Qy 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Db 182 LRAAEYSNMVFLGKGNSSKKQLVQVEAGAIGFKLHEDWGTTPSAIDHCLSVADRYD 241
Qy 242 VQVCHTDTVNEAGYVDDTLNANWGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSITP 301
Db 242 VQVCHTDTVNEAGYVDDTLNANWGRAIHAYHIEGAGGSHSPDVTMAGELNLPSSITP 301
Qy 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIIRPGSAEEDVLHDMGVIAWTSDD 361
Db 302 TIPTYTINTVAEHLDMMLTCHLDKRIREDLOFSQSRIKRGSIIRPGSAEEDVLHDMGVIAWTSDD 361
Qy 362 SQAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Db 362 SQAMGRAGEVIRPTWQTADKNKKEFKLPEDGKNDNFRIKRYISKYITINPALTHGVSEY 421
Qy 422 IGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 481
Db 422 IGSVEEGKIADLVVWNPAPFVGKPKIVIKGMVVFSEMGDSNASVPTPQPVYVREMFHGH 481
Qy 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
Db 482 GKAKEDTSTTFVSKVAYENGVEKGLERQVLPVKNCRNITKKDFKNDKTAKITVDPKT 541
Qy 542 FEVFDGKLCIT 552
Db 542 FEVFDGKLCIT 552

RESULT 15

US-09-252-991A-26887
; Sequence 26887, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26887
; LENGTH: 566
; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26887

Query Match		61.0%;	Score 1828;	DB 2;	Length 566;
Best Local Similarity		59.2%;	Pred. No. 3.5e-168;		
Matches	336;	Conservative	57;	Mismatches	143;
				Indels	2;
				Gaps	1;
QY	1	MMKKQEVYNTYVGP	TKDKVRLGDTDLMAEVEHDYTYT	GEEELKPGAGKTIREGMGQSN	60
DB	1	MMKISQAVADMFGP	TVGDRVLADTDLMTIEVERDFTV	GEEVKFGGKVRDGMGQSLG	60
QY	61	DENTLDLVTNAMI	TDYGIKADIGIKNGKIHGICAGK	KDMQDCVSPHMVVGVGTEAL	120
DB	61	AAQVVDVTITNAL	ILDHGWKADVGLKDGRIQAI	GKAGNPDIQGV--NIAIGAGTEVI	118
QY	121	AGEGMIITAGGID	STHFLSPQOFTALANGVTPV	GGTGPVDGTNATTITPGKWNLHR	180
DB	119	AGEGMILTAGGID	THIFICPOQIEALMSGVT	TWIGGTGPAAGTNATCTSGPWHMAR	178
QY	181	MLRAAEYSMN	VFLGKGNSSKKQLVEQVEAGAIG	FKLHEDWGTTPSAIDHCLSVADEY	240
DB	179	MLQAADAFPMN	IGFTGKNASLPLEEQVLAGAIG	LKLHEDWGSTPAIDNCLVEAERH	238
QY	241	DVQVCIHDTVNE	AGYVDDTLNAMNGRAIHAYHIE	GAGGCHSPDVTMAGELNLPSTT	300
DB	239	DIQVAIHDTDLN	ESGFVETTLGAFKRTIHTYHTE	GAGGCHAPDIIKACGFANVLPSTN	298
QY	301	PTIPYTIINTVA	EHLDMLTCHHLDKRIREDLO	FOSRIRPGSIAAEDVLHDMGVIA	360
DB	299	PTRPFRNTIDB	HLDMVCHHLDPALAEADVAF	ESRIIRRETTAAEDILHDLGAF	358
QY	361	DSQAMGRAGEVI	PRWTQADKNKBFGLPEDGK	DNDFRIKRYISKYTIINPALTHG	420
DB	359	DSQAMGRVGEVI	TRTWQTADKMKQGRLDG	DGARNDFRARRYIAKYTIINPAI	418
QY	421	YIGSVEEGKIAD	LVVNPAPFGVKPKIVIKGM	VFSEMGDSNASVPTPQPVYIR	480
DB	419	EVGSVEAGKWAD	LWLWRPAPFGVKPSLILK	GGAIAASLMGDIINGSIPTQPV	478
QY	481	HGKAKEDTSIT	FVSKVAYENGVEKGLERQ	VLVKNCHNITKKOPKFNDKTA	540
DB	479	YAGSRHATSLT	FVSOAAFAAGVPQQLGLR	KAIGWVSGCRGVQKTDLIHNG	538
QY	541	TPEVFVDGKLC	TSKPTSOVPLAQR	YTF 568	
DB	539	NYQVRADGQL	LWCEPADVLPMAQR	YFLP 566	

Search completed: November 28, 2005, 08:25:42
Job time : 48.4987 secs